

# (12) UK Patent Application (19) GB (11) 2 372 993 (13) A

(43) Date of A Publication 11.09.2002

(21) Application No 0126380.5

(22) Date of Filing 02.11.2001

(30) Priority Data

(31) 0026951 (32) 03.11.2000 (33) GB  
(31) 0109787 (32) 20.04.2001

(71) Applicant(s)

**SmithKline Beecham p.l.c.**  
(Incorporated in the United Kingdom)  
New Horizons Court, Great West Road, BRENTFORD,  
Middlesex, TW8 9EP, United Kingdom

**SmithKline Beecham Corporation**  
(Incorporated in USA - Pennsylvania)  
One Franklin Plaza, P.O. Box 7927, Philadelphia,  
Pennsylvania 19103, United States of America

(72) Inventor(s)

**John Beresford Davis**  
**Martin Gunthorpe**  
**Philip David Hayes**  
**Rosemary Elizabeth Kelsell**

(51) INT CL<sup>7</sup>

**C07K 14/705 , A61P 13/00 25/00 29/02 37/02 , C12N  
15/63 , G01N 33/53**

(52) UK CL (Edition T )

**C3H HB7P H656 H686  
G1B BAX  
U1S S2411 S2415 S2416 S2417 S2418**

(56) Documents Cited

**WO 2002/012340 A WO 2002/000722 A**

(58) Field of Search

**ONLINE: WPI, EPODOC, JAPIO, MEDLINE, EMBASE,  
BIOSIS, SCISEARCH, CAPLUS, BLASTp, BLASTn**

(74) Agent and/or Address for Service

**GlaxoSmithKline**  
**Corporate Intellectual Property, Two New Horizons**  
**Court, BRENTFORD, Middlesex, TW8 9EP,**  
**United Kingdom**

(54) Abstract Title

**Vanilloid Receptor 6**

(57) Vanilloid receptor 6 (VANILREP6) polypeptides and polynucleotides and methods for producing these ion channels by recombinant techniques are disclosed. Also disclosed are methods for utilising VANILREP6 polynucleotides and polypeptides in pharmacological assays. Splice variants of VANILREP6 are also disclosed. Such peptides may be of use in the treatment of disorders including pain, neurological disorders, inflammatory disorders, ischemia, neurodegeneration and inflammatory bowel disease.

GB 2 372 993 A

## Novel Compounds

### Field of the Invention

This invention relates to newly identified polypeptides and polynucleotides encoding such polypeptides, to their use in diagnosis and in identifying compounds that may be agonists, antagonists that are potentially useful in therapy, and to production of such polypeptides and polynucleotides.

### Background of the Invention

The drug discovery process is currently undergoing a fundamental revolution as it embraces "functional genomics", that is, high throughput genome- or gene-based biology. This approach as a means to identify genes and gene products as therapeutic targets is rapidly superseding earlier approaches based on "positional cloning". A phenotype, that is a biological function or genetic disease, would be identified and this would then be tracked back to the responsible gene, based on its genetic map position.

Functional genomics relies heavily on high-throughput DNA sequencing technologies and the various tools of bioinformatics to identify gene sequences of potential interest from the many molecular biology databases now available. There is a continuing need to identify and characterise further genes and their related polypeptides/proteins, as targets for drug discovery.

### Summary of the Invention

The present invention relates to VANILREP6, in particular VANILREP6 polypeptides and VANILREP6 polynucleotides, recombinant materials and methods for their production. Such polypeptides and polynucleotides are of interest in relation to methods of treatment of certain diseases, including, but not limited to, pain, chronic pain, neuropathic pain, postoperative pain, rheumatoid arthritic pain, neuralgia, migraine, epilepsy, visceral pain, cystitis, irritable bowel syndrome, neuropathies, algesia, motion sickness, balance disorders, nerve injury, ischaemia, neurodegeneration, stroke, incontinence, asthma and inflammatory disorders,, hereinafter referred to as "diseases of the invention". In a further aspect, the invention relates to methods for identifying agonists and antagonists (*e.g.*, inhibitors) using the materials provided by the invention, and treating conditions associated with VANILREP6 imbalance with the identified compounds. In a still further aspect, the invention relates to diagnostic assays for detecting diseases associated with inappropriate VANILREP6 activity or levels.

## Description of the Invention

In a first aspect, the present invention relates to VANILREP6 polypeptides. Such polypeptides include:

- (a) an isolated polypeptide encoded by a polynucleotide comprising the sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9;
- (b) an isolated polypeptide comprising a polypeptide sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the polypeptide sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10;
- (c) an isolated polypeptide comprising the polypeptide sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10;
- (d) an isolated polypeptide having at least 95%, 96%, 97%, 98%, or 99% identity to the polypeptide sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10;
- (e) the polypeptide sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10; and
- (f) an isolated polypeptide having or comprising a polypeptide sequence that has an Identity Index of 0.95, 0.96, 0.97, 0.98, or 0.99 compared to the polypeptide sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10;
- (g) fragments and variants of such polypeptides in (a) to (f).

Polypeptides of the present invention are believed to be members of the Ion channel family of polypeptides. They are therefore of interest because they are related to the VR1 channel which is associated with the mechanism of action of capsaicin (a vanilloid compound), a constituent of chilli peppers. Capsaicin elicits a sensation of burning pain by selectively activating sensory neurons that convey information about noxious stimuli to the central nervous system. The channels are permeable to cations and exhibit a notable preference for divalent cations, particularly calcium ions. The level of calcium ion permeability exceeds that observed for most non-selective cation channels and is similar to values observed for NMDA-type glutamate receptors and  $\alpha 7$  nicotinic acetylcholine receptors, both of which are noted for this property. Ion channels are particularly important in homeostasis and signalling pathways, thus being attractive targets for therapeutic intervention. The biological properties of the VANILREP6 are hereinafter referred to as "biological activity of VANILREP6" or "VANILREP6 activity". Preferably, a polypeptide of the present invention exhibits at least one biological activity of VANILREP6

Polypeptides of the present invention also includes variants of the aforementioned polypeptides, including all allelic forms and splice variants. Such polypeptides vary from the reference polypeptide by insertions, deletions, and substitutions that may be conservative or non-conservative, or any combination thereof. Particularly preferred variants are those in which several,

for instance from 50 to 30, from 30 to 20, from 20 to 10, from 10 to 5, from 5 to 3, from 3 to 2, from 2 to 1 or 1 amino acids are inserted, substituted, or deleted, in any combination.

Preferred fragments of polypeptides of the present invention include an isolated polypeptide comprising an amino acid sequence having at least 30, 50 or 100 contiguous amino acids from the amino acid sequence of SEQ ID NO: 2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10, or an isolated polypeptide comprising an amino acid sequence having at least 30, 50 or 100 contiguous amino acids truncated or deleted from the amino acid sequence of SEQ ID NO: 2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10. Preferred fragments are biologically active fragments that mediate the biological activity of VANILREP6, including those with a similar activity or an improved activity, or with a decreased undesirable activity. Also preferred are those fragments that are antigenic or immunogenic in an animal, especially in a human.

Fragments of the polypeptides of the invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, these variants may be employed as intermediates for producing the full-length polypeptides of the invention. The polypeptides of the present invention may be in the form of the "mature" protein or may be a part of a larger protein such as a precursor or a fusion protein. It is often advantageous to include an additional amino acid sequence that contains secretory or leader sequences, pro-sequences, sequences that aid in purification, for instance multiple histidine residues, or an additional sequence for stability during recombinant production.

Polypeptides of the present invention can be prepared in any suitable manner, for instance by isolation from naturally occurring sources, from genetically engineered host cells comprising expression systems (*vide infra*) or by chemical synthesis, using for instance automated peptide synthesizers, or a combination of such methods. Means for preparing such polypeptides are well understood in the art.

In a further aspect, the present invention relates to VANILREP6 polynucleotides. Such polynucleotides include:

- (a) an isolated polynucleotide comprising a polynucleotide sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the polynucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9;
- (b) an isolated polynucleotide comprising the polynucleotide of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9;
- (c) an isolated polynucleotide having at least 95%, 96%, 97%, 98%, or 99% identity to the polynucleotide of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9;
- (d) the isolated polynucleotide of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9;

- (e) an isolated polynucleotide comprising a polynucleotide sequence encoding a polypeptide sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the polypeptide sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10;
- (f) an isolated polynucleotide comprising a polynucleotide sequence encoding the polypeptide of  
5 SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10;
- (g) an isolated polynucleotide having a polynucleotide sequence encoding a polypeptide sequence having at least 95%, 96%, 97%, 98%, or 99% identity to the polypeptide sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10;
- (h) an isolated polynucleotide encoding the polypeptide of SEQ ID NO:2, SEQ ID NO:4, SEQ ID  
10 NO:6 or SEQ ID NO:10;
- (i) an isolated polynucleotide having or comprising a polynucleotide sequence that has an Identity Index of 0.95, 0.96, 0.97, 0.98, or 0.99 compared to the polynucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9;
- (j) an isolated polynucleotide having or comprising a polynucleotide sequence encoding a  
15 polypeptide sequence that has an Identity Index of 0.95, 0.96, 0.97, 0.98, or 0.99 compared to the polypeptide sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10; and polynucleotides that are fragments and variants of the above mentioned polynucleotides or that are complementary to above mentioned polynucleotides, over the entire length thereof.

Preferred fragments of polynucleotides of the present invention include an isolated  
20 polynucleotide comprising an nucleotide sequence having at least 15, 30, 50 or 100 contiguous nucleotides from the sequence of SEQ ID NO: 1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9, or an isolated polynucleotide comprising an sequence having at least 30, 50 or 100 contiguous nucleotides truncated or deleted from the sequence of SEQ ID NO: 1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9.

25 Preferred variants of polynucleotides of the present invention include splice variants, allelic variants, and polymorphisms, including polynucleotides having one or more single nucleotide polymorphisms (SNPs).

Polynucleotides of the present invention also include polynucleotides encoding polypeptide variants that comprise the amino acid sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or  
30 SEQ ID NO:10 and in which several, for instance from 50 to 30, from 30 to 20, from 20 to 10, from 10 to 5, from 5 to 3, from 3 to 2, from 2 to 1 or 1 amino acid residues are substituted, deleted or added, in any combination.

In a further aspect, the present invention provides polynucleotides that are RNA transcripts of the DNA sequences of the present invention. Accordingly, there is provided an  
35 RNA polynucleotide that:

(a) comprises an RNA transcript of the DNA sequence encoding the polypeptide of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10;

(b) is the RNA transcript of the DNA sequence encoding the polypeptide of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10;

5 (c) comprises an RNA transcript of the DNA sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9; or

(d) is the RNA transcript of the DNA sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9;

and RNA polynucleotides that are complementary thereto.

10 The polynucleotide sequences of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 and SEQ ID NO:9 show homology with the VR1 nonselective cation channel (Hayes et al., Pain. 2000 Nov 1;88(2):205-215).

The polynucleotide sequence of SEQ ID NO:1 is a cDNA sequence that encodes VANILREP6 polypeptides. The DNA sequence of SEQ ID NO:1 includes a number of  
15 polymorphic variants as described more fully in Table 1 (the numbering of nucleotides in Table 1 follows that of SEQ ID NO:5). One polynucleotide of SEQ ID NO:1 encodes the polypeptide of SEQ ID NO:2. The polynucleotide sequence encoding the polypeptide of SEQ ID NO:2 may be identical to the polypeptide encoding sequence of SEQ ID NO:1 or it may be a sequence other than SEQ ID NO:1, which, as a result of the redundancy (degeneracy) of the genetic code, also  
20 encodes the polypeptide of SEQ ID NO:2. The polypeptide of the SEQ ID NO:2 is related to other proteins of the Ion channel family, having homology and/or structural similarity with the VR1 nonselective cation channel (Hayes et al., Pain. 2000 Nov 1;88(2):205-215).

Splice variants of the VANILREP6 polynucleotides, and the polypeptides encoded by them also form part of the present invention. In one preferred embodiment the splice variant is that shown  
25 as the polynucleotide of SEQ ID NO:3 which has a 59bp sequence deletion compared to the polynucleotide of SEQ ID NO:1. The DNA sequence of SEQ ID NO:3 includes a number of polymorphic variants as described more fully hereinabove and in Table 1. One polynucleotide of SEQ ID NO:3 encodes the polypeptide of SEQ ID NO:4.

In a further preferred embodiment the splice variant is that shown as the polynucleotide  
30 of SEQ ID NO:5. An alignment of the sequences of SEQ ID NO:5 and SEQ ID NO:7 (5' untranslated region upstream of the ATG start codon of SEQ ID NO:1) shows common sequence downstream (3' to) nucleotide 70 of SEQ ID NO:5 and nucleotide 146 of SEQ ID NO:7. However the alignment shows that the sequences upstream (5' to) the aforementioned nucleotides display little homology with each other. In addition, SEQ ID NO 5 also has a  
35 deletion for the triplet CAG (at position 2078 to 2080 of SEQ ID NO:1) as a result of

alternative splicing. This results in a polypeptide of 790 amino acids shown as SEQ ID NO:6. The DNA sequence of SEQ ID NO:5 includes a number of polymorphic variants as described more fully in Table 1. One polynucleotide of SEQ ID NO:5 encodes the polypeptide of SEQ ID NO:6.

In a further preferred embodiment the splice variant is that shown as the polynucleotide of SEQ ID NO:9. The polynucleotide of SEQ ID NO:9 contains the same splice variations as SEQ ID NO:5 but in addition, has a further 87 bp deletion. The DNA sequence of SEQ ID NO:9 includes a number of polymorphic variants as described more fully in Table 1. One polynucleotide of SEQ ID NO:9 encodes the polypeptide of SEQ ID NO:10.

Polymorphic variants of the polynucleotides of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9, which may or may not lead to changes in the encoded polypeptides (for example those of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10, including, but not limited to, those shown in Table 1 also form part of the present invention.

Table 1: Examples of VANILREP6 polymorphic variants. Nucleotide numbering is based on SEQ ID NO:5 where base no. 1 is the first base of the ATG start codon. Amino acid numbering is derived from SEQ ID NO:6.

Polymorphism	Effect on encoded polypeptide
G(270)A (shown as "R" in SEQ ID NO:5)	Silent
A(349)G (shown as "R" in SEQ ID NO:5)	amino acid R(117)G substitution
A(558)C (shown as "M" in SEQ ID NO:5)	Silent
G(936)A (shown as "R" in SEQ ID NO:5)	Silent
C(1746)T (shown as "Y" in SEQ ID NO:5)	Silent
C(1878)T (shown as "Y" in SEQ ID NO:5)	Silent
C(1923)T (shown as "Y" in SEQ ID NO:5)	Silent

The nucleotide sequence of SEQ ID NO:8 is a cDNA sequence and comprises a polypeptide encoding sequence (nucleotide 169 to 2541 which is equivalent to the coding sequence of SEQ ID NO:5). Additional 5' sequence obtained from RACE clones has been added on to generate this SEQ ID NO:8. Exon 1 (1-166 bp) Exon 2 (167-287 bp), Exon 3 (288-392 bp), Exon 4 (393-479 bp), Exon 5 (480-634 bp), Exon 6 (635-811 bp), Exon 7 (812-952 bp), Exon 8 (953-1233 bp), Exon 9 (1234-1410 bp), Exon 10 (1411-1569 bp), Exon 11 (1570-1671 bp), Exon 12 (1672-1745 bp), Exon 13 (1746-1911 bp), Exon 14 (1912-1978 bp), Exon

15 (1979-2253 bp), Exon 16 (2254-2366 bp), Exon 17 (2367-2446 bp), Exon 18 (2447-2612 bp) encodes a polypeptide of 790 amino acids, the polypeptide of SEQ ID NO:6.

Knowledge of the exon-intron structure of VANILREP6 can be used for mutation screening, for example as a diagnostic test for diseases which may be caused by alterations of VANILREP6 . The screening of genomic DNA is desirable for the analysis of non-coding regions, such as upstream regulatory elements and intron splice sites. It is also useful for cases where mRNA is not readily available for mutation analysis. It will be important to determine the frequencies of the aforementioned polymorphisms in the general population and to ascertain whether any of these are indeed associated with disease. The genomic structure is also useful in analysing the splice variants of VANILREP6. For example, SEQ ID NO:3 represents a splice variant of SEQ ID NO:1, missing the first 59 bp of exon 13 as a result of splicing on to a cryptic splice acceptor site within this exon. SEQ ID NO:5 contains two splice variations. It diverges from SEQ ID NO:1 within exon 2, and the (CAG) deletion after 2278 bp (nucleotides 2078 to 2080 of SEQ ID NO:1) occurs by splicing to a cryptic splice acceptor site 3 bp into exon 18. SEQ ID NO:9 contains the same variations as SEQ ID NO:5, but in addition, is deleted for exon 4. Splice variants are important because they may have different functions and different expression patterns. Knowledge of the genomic structure is also important for the generation of animal models. Such models may be used to study the function of VANILREP6 and for drug screening studies. For example, mouse knock-out models typically have a selection marker, which upon insertion into a coding exon, ablate the functioning of the targeted allele.

Preferred polypeptides and polynucleotides of the present invention are expected to have, *inter alia*, similar biological functions/properties to their homologous polypeptides and polynucleotides. Furthermore, preferred polypeptides and polynucleotides of the present invention have at least one VANILREP6 activity.

The VANILREP6 polynucleotides of the present invention may be obtained using standard cloning and screening techniques from a cDNA library derived from mRNA in cells of human, for example, whole brain, corpus callosum, testis, colon, colorectal adenocarcinoma, small intestine, fetal small intestine and bladder. The splice variant polynucleotide (SEQ ID NO:3) may be obtained from, for example, corpus callosum, hippocampus, heart and cerebellum; the splice variant polynucleotide (SEQ ID NO:5) may be obtained from, for example, small intestine, fetal small intestine, colon, colorectal adenocarcinoma and bladder; and the splice variant polynucleotide (SEQ ID NO:9) may be obtained from, for example, small intestine, fetal small intestine and colon (see for instance, Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989)). Polynucleotides of



the invention can also be obtained from natural sources such as genomic DNA libraries or can be synthesized using well known and commercially available techniques.

When polynucleotides of the present invention are used for the recombinant production of polypeptides of the present invention, the polynucleotide may include the coding sequence for the mature polypeptide, by itself, or the coding sequence for the mature polypeptide in reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, or pro- or prepro- protein sequence, or other fusion peptide portions. For example, a marker sequence that facilitates purification of the fused polypeptide can be encoded. In certain preferred embodiments of this aspect of the invention, the marker sequence is a hexa-histidine peptide, as provided in the pQE vector (Qiagen, Inc.) and described in Gentz *et al.*, Proc Natl Acad Sci USA (1989) 86:821-824, or is an HA tag. The polynucleotide may also contain non-coding 5' and 3' sequences, such as transcribed, non-translated sequences, splicing and polyadenylation signals, ribosome binding sites and sequences that stabilize mRNA.

Polynucleotides that are identical, or have sufficient identity to a polynucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9 may be used as hybridization probes for cDNA and genomic DNA or as primers for a nucleic acid amplification reaction (for instance, PCR). Such probes and primers may be used to isolate full-length cDNAs and genomic clones encoding polypeptides of the present invention and to isolate cDNA and genomic clones of other genes (including genes encoding paralogs from human sources and orthologs and paralogs from species other than human that have a high sequence similarity to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9 typically at least 95% identity. Preferred probes and primers will generally comprise at least 15 nucleotides, preferably, at least 30 nucleotides and may have at least 50, if not at least 100 nucleotides. Particularly preferred probes will have between 30 and 50 nucleotides. Particularly preferred primers will have between 20 and 25 nucleotides.

A polynucleotide encoding a polypeptide of the present invention, including homologs from species other than human may be obtained by a process comprising the steps of screening a library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9 or a fragment thereof, preferably of at least 15 nucleotides; and isolating full-length cDNA and genomic clones containing said polynucleotide sequence. Such hybridization techniques are well known to the skilled artisan. Preferred stringent hybridization conditions include overnight incubation at 42°C in a solution comprising: 50% formamide, 5xSSC (150mM NaCl, 15mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10 % dextran sulfate, and 20 microgram/ml denatured, sheared salmon sperm DNA; followed by washing the filters in 0.1x SSC at about 65°C. Thus the present invention also includes isolated polynucleotides, preferably with a nucleotide sequence of at least 100, obtained by screening a library under stringent hybridization conditions with a labeled probe having the

sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9 or a fragment thereof, preferably of at least 15 nucleotides.

The skilled artisan will appreciate that, in many cases, an isolated cDNA sequence will be incomplete, in that the region coding for the polypeptide does not extend all the way through to the 5' terminus. This is a consequence of reverse transcriptase, an enzyme with inherently low "processivity" (a measure of the ability of the enzyme to remain attached to the template during the polymerisation reaction), failing to complete a DNA copy of the mRNA template during first strand cDNA synthesis.

There are several methods available and well known to those skilled in the art to obtain full-length cDNAs, or extend short cDNAs, for example those based on the method of Rapid Amplification of cDNA ends (RACE) (see, for example, Frohman et al., Proc Nat Acad Sci USA 85, 8998-9002, 1988). Recent modifications of the technique, exemplified by the Marathon (trade mark) technology (Clontech Laboratories Inc.) for example, have significantly simplified the search for longer cDNAs. In the Marathon (trade mark) technology, cDNAs have been prepared from mRNA extracted from a chosen tissue and an 'adaptor' sequence ligated onto each end. Nucleic acid amplification (PCR) is then carried out to amplify the "missing" 5' end of the cDNA using a combination of gene specific and adaptor specific oligonucleotide primers. The PCR reaction is then repeated using 'nested' primers, that is, primers designed to anneal within the amplified product (typically an adapter specific primer that anneals further 3' in the adaptor sequence and a gene specific primer that anneals further 5' in the known gene sequence). The products of this reaction can then be analyzed by DNA sequencing and a full-length cDNA constructed either by joining the product directly to the existing cDNA to give a complete sequence, or carrying out a separate full-length PCR using the new sequence information for the design of the 5' primer.

Recombinant polypeptides of the present invention may be prepared by processes well known in the art from genetically engineered host cells comprising expression systems. Accordingly, in a further aspect, the present invention relates to expression systems comprising a polynucleotide or polynucleotides of the present invention, to host cells which are genetically engineered with such expression systems and to the production of polypeptides of the invention by recombinant techniques. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention.

For recombinant production, host cells can be genetically engineered to incorporate expression systems or portions thereof for polynucleotides of the present invention. Polynucleotides may be introduced into host cells by methods described in many standard laboratory manuals, such as Davis et al., Basic Methods in Molecular Biology (1986) and Sambrook *et al.*(*ibid*). Preferred methods of introducing polynucleotides into host cells include, for instance, calcium phosphate

transfection, DEAE-dextran mediated transfection, transvection, micro-injection, cationic lipid-mediated transfection, electroporation, transduction, scrape loading, ballistic introduction or infection.

Representative examples of appropriate hosts include bacterial cells, such as *Streptococci*,  
5 *Staphylococci*, *E. coli*, *Streptomyces* and *Bacillus subtilis* cells; fungal cells, such as yeast cells and *Aspergillus* cells; insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, HeLa, C127, 3T3, BHK, HEK 293 and Bowes melanoma cells; and plant cells.

A great variety of expression systems can be used, for instance, chromosomal, episomal and virus-derived systems, *e.g.*, vectors derived from bacterial plasmids, from bacteriophage, from  
10 transposons, from yeast episomes, from insertion elements, from yeast chromosomal elements, from viruses such as baculoviruses, papova viruses, such as SV40, vaccinia viruses, adenoviruses, fowl pox viruses, pseudorabies viruses and retroviruses, and vectors derived from combinations thereof, such as those derived from plasmid and bacteriophage genetic elements, such as cosmids and phagemids. The expression systems may contain control regions that regulate as well as engender  
15 expression. Generally, any system or vector that is able to maintain, propagate or express a polynucleotide to produce a polypeptide in a host may be used. The appropriate polynucleotide sequence may be inserted into an expression system by any of a variety of well-known and routine techniques, such as, for example, those set forth in Sambrook *et al.*, (*ibid*). Appropriate secretion signals may be incorporated into the desired polypeptide to allow secretion of the translated protein  
20 into the lumen of the endoplasmic reticulum, the periplasmic space or the extracellular environment. These signals may be endogenous to the polypeptide or they may be heterologous signals.

If a polypeptide of the present invention is to be expressed for use in screening assays, it is generally preferred that the polypeptide be produced at the surface of the cell. In this event, the cells may be harvested prior to use in the screening assay. If the polypeptide is secreted into the  
25 medium, the medium can be recovered in order to recover and purify the polypeptide. If produced intracellularly, the cells must first be lysed before the polypeptide is recovered.

Polypeptides of the present invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography,  
30 hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during intracellular synthesis, isolation and/or purification.

35 Polynucleotides of the present invention may be used as diagnostic reagents, through detecting mutations in the associated gene. Detection of a mutated form of the gene characterized by

the polynucleotide of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9 in the cDNA or genomic sequence and which is associated with a dysfunction will provide a diagnostic tool that can add to, or define, a diagnosis of a disease, or susceptibility to a disease, which results from under-expression, over-expression or altered spatial or temporal expression of the gene. Individuals  
5 carrying mutations in the gene may be detected at the DNA level by a variety of techniques well known in the art.

Nucleic acids for diagnosis may be obtained from a subject's cells, such as from blood, urine, saliva, tissue biopsy or autopsy material. The genomic DNA may be used directly for detection or it may be amplified enzymatically by using PCR, preferably RT-PCR, or other amplification  
10 techniques prior to analysis. RNA or cDNA may also be used in similar fashion. Deletions and insertions can be detected by a change in size of the amplified product in comparison to the normal genotype. Point mutations can be identified by hybridizing amplified DNA to labeled VANILREP6 nucleotide sequences. Perfectly matched sequences can be distinguished from mismatched duplexes by RNase digestion or by differences in melting temperatures. DNA sequence difference may also  
15 be detected by alterations in the electrophoretic mobility of DNA fragments in gels, with or without denaturing agents, or by direct DNA sequencing (see, for instance, Myers *et al.*, Science (1985) 230:1242). Sequence changes at specific locations may also be revealed by nuclease protection assays, such as RNase and S1 protection or the chemical cleavage method (see Cotton *et al.*, Proc Natl Acad Sci USA (1985) 85: 4397-4401).

An array of oligonucleotides probes comprising VANILREP6 polynucleotide sequence or fragments thereof can be constructed to conduct efficient screening of *e.g.*, genetic mutations. Such arrays are preferably high density arrays or grids. Array technology methods are well known and have general applicability and can be used to address a variety of questions in molecular genetics including gene expression, genetic linkage, and genetic variability, see, for example, M. Chee *et al.*,  
20 Science, 274, 610-613 (1996) and other references cited therein.

Detection of abnormally decreased or increased levels of polypeptide or mRNA expression may also be used for diagnosing or determining susceptibility of a subject to a disease of the invention. Decreased or increased expression can be measured at the RNA level using any of the methods well known in the art for the quantitation of polynucleotides, such as, for example,  
30 nucleic acid amplification, for instance PCR, RT-PCR, RNase protection, Northern blotting and other hybridization methods. Assay techniques that can be used to determine levels of a protein, such as a polypeptide of the present invention, in a sample derived from a host are well-known to those of skill in the art. Such assay methods include radio-immunoassays, competitive-binding assays, Western Blot analysis and ELISA assays.

35 Thus in another aspect, the present invention relates to a diagnostic kit comprising:

- (a) a polynucleotide of the present invention, preferably the nucleotide sequence of SEQ ID NO: 1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9 or a fragment or an RNA transcript thereof;
- (b) a nucleotide sequence complementary to that of (a);
- (c) a polypeptide of the present invention, preferably the polypeptide of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10 or a fragment thereof; or
- (d) an antibody to a polypeptide of the present invention, preferably to the polypeptide of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component. Such a kit will be of use in diagnosing a disease or susceptibility to a disease, particularly diseases of the invention, amongst others.

The polynucleotide sequences of the present invention are valuable for chromosome localisation studies. The sequence is specifically targeted to, and can hybridize with, a particular location on an individual human chromosome. The mapping of relevant sequences to chromosomes according to the present invention is an important first step in correlating those sequences with gene associated disease. Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. Such data are found in, for example, V. McKusick, Mendelian Inheritance in Man (available on-line through Johns Hopkins University Welch Medical Library). The relationship between genes and diseases that have been mapped to the same chromosomal region are then identified through linkage analysis (co-inheritance of physically adjacent genes). Precise human chromosomal localisations for a genomic sequence (gene fragment etc.) can be determined using Radiation Hybrid (RH) Mapping (Walter, M. Spillett, D., Thomas, P., Weissenbach, J., and Goodfellow, P., (1994) A method for constructing radiation hybrid maps of whole genomes, *Nature Genetics* 7, 22-28). A number of RH panels are available from Research Genetics (Huntsville, AL, USA) e.g. the GeneBridge4 RH panel (*Hum Mol Genet* 1996 Mar;5(3):339-46 A radiation hybrid map of the human genome. Gyapay G, Schmitt K, Fizames C, Jones H, Vega-Czarny N, Spillett D, Muselet D, Prud'Homme JF, Dib C, Auffray C, Morissette J, Weissenbach J, Goodfellow PN). To determine the chromosomal location of a gene using this panel, 93 PCRs are performed using primers designed from the gene of interest on RH DNAs. Each of these DNAs contains random human genomic fragments maintained in a hamster background (human / hamster hybrid cell lines). These PCRs result in 93 scores indicating the presence or absence of the PCR product of the gene of interest. These scores are compared with scores created using PCR products from genomic sequences of known location. This comparison is conducted at <http://www.genome.wi.mit.edu/>. The gene of the present invention maps to human chromosome 17p13. According to the available genomic sequences (AC027796, version 4) it is situated less than 10 kb away from VANILREP1, and is transcribed in the same direction.

The polynucleotide sequences of the present invention are also valuable tools for tissue expression studies. Such studies allow the determination of expression patterns of polynucleotides of the present invention which may give an indication as to the expression patterns of the encoded polypeptides in tissues, by detecting the mRNAs that encode them. The techniques used are well known in the art and include in situ hybridization techniques to clones arrayed on a grid, such as cDNA microarray hybridization (Schena *et al.*, Science, 270, 467-470, 1995 and Shalon *et al.*, Genome Res, 6, 639-645, 1996) and nucleotide amplification techniques such as PCR. A preferred method uses the TAQMAN (Trade mark) technology available from Perkin Elmer. Results from these studies can provide an indication of the normal function of the polypeptide in the organism. In addition, comparative studies of the normal expression pattern of mRNAs with that of mRNAs encoded by an alternative form of the same gene (for example, one having an alteration in polypeptide coding potential or a regulatory mutation) can provide valuable insights into the role of the polypeptides of the present invention, or that of inappropriate expression thereof in disease. Such inappropriate expression may be of a temporal, spatial or simply quantitative nature.

The polynucleotides of the present invention are expressed in the nervous system and to a lesser extent in peripheral tissues, including pituitary, heart, skeletal muscle, stomach, intestine and placenta. Expression across regions of the central nervous system is uniform, including expression in thalamus, cortex, hippocampus, hypothalamus, corpus callosum, spinal cord, amygdala, caudate nucleus and putamen. Expression in the dorsal root ganglia is three fold that of whole brain.

A further aspect of the present invention relates to antibodies. The polypeptides of the invention or their fragments, or cells expressing them, can be used as immunogens to produce antibodies that are immunospecific for polypeptides of the present invention. The term "immunospecific" means that the antibodies have substantially greater affinity for the polypeptides of the invention than their affinity for other related polypeptides in the prior art.

Antibodies generated against polypeptides of the present invention may be obtained by administering the polypeptides or epitope-bearing fragments, or cells to an animal, preferably a non-human animal, using routine protocols. For preparation of monoclonal antibodies, any technique which provides antibodies produced by continuous cell line cultures can be used. Examples include the hybridoma technique (Kohler, G. and Milstein, C., Nature (1975) 256:495-497), the trioma technique, the human B-cell hybridoma technique (Kozbor *et al.*, Immunology Today (1983) 4:72) and the EBV-hybridoma technique (Cole *et al.*, Monoclonal Antibodies and Cancer Therapy, 77-96, Alan R. Liss, Inc., 1985).

Techniques for the production of single chain antibodies, such as those described in U.S. Patent No. 4,946,778, can also be adapted to produce single chain antibodies to polypeptides of this

invention. Also, transgenic mice, or other organisms, including other mammals, may be used to express humanized antibodies.

The above-described antibodies may be employed to isolate or to identify clones expressing the polypeptide or to purify the polypeptides by affinity chromatography. Antibodies against polypeptides of the present invention may also be employed to treat diseases of the invention, amongst others.

Polypeptides and polynucleotides of the present invention may also be used as vaccines. Accordingly, in a further aspect, the present invention relates to a method for inducing an immunological response in a mammal that comprises inoculating the mammal with a polypeptide of the present invention, adequate to produce antibody and/or T cell immune response, including, for example, cytokine-producing T cells or cytotoxic T cells, to protect said animal from disease whether that disease is already established within the individual or not. An immunological response in a mammal may also be induced by a method comprises delivering a polypeptide of the present invention *via* a vector directing expression of the polynucleotide and coding for the polypeptide *in vivo* in order to induce such an immunological response to produce antibody to protect said animal from diseases of the invention. One way of administering the vector is by accelerating it into the desired cells as a coating on particles or otherwise. Such nucleic acid vector may comprise DNA, RNA, a modified nucleic acid, or a DNA/RNA hybrid. For use as a vaccine, a polypeptide or a nucleic acid vector will be normally provided as a vaccine formulation (composition). The formulation may further comprise a suitable carrier. Since a polypeptide may be broken down in the stomach, it is preferably administered parenterally (for instance, subcutaneous, intra-muscular, intravenous, or intra-dermal injection). Formulations suitable for parenteral administration include aqueous and non-aqueous sterile injection solutions that may contain anti-oxidants, buffers, bacteriostats and solutes that render the formulation isotonic with the blood of the recipient; and aqueous and non-aqueous sterile suspensions that may include suspending agents or thickening agents. The formulations may be presented in unit-dose or multi-dose containers, for example, sealed ampoules and vials and may be stored in a freeze-dried condition requiring only the addition of the sterile liquid carrier immediately prior to use. The vaccine formulation may also include adjuvant systems for enhancing the immunogenicity of the formulation, such as oil-in water systems and other systems known in the art. The dosage will depend on the specific activity of the vaccine and can be readily determined by routine experimentation.

Polypeptides of the present invention have one or more biological functions that are of relevance in one or more disease states, in particular the diseases of the invention hereinbefore mentioned. It is therefore useful to identify pharmacological or biophysical methods, such as

increased temperature, that stimulate or inhibit the function or level of the polypeptide. Accordingly, in a further aspect, the present invention provides for a method of screening compounds to identify those that stimulate or inhibit the function or level of the polypeptide. Such methods identify agonists or antagonists that may be employed for therapeutic and prophylactic purposes for such diseases of the invention as hereinbefore mentioned. Compounds may be identified from a variety of sources, for example, cells, cell-free preparations, chemical libraries, collections of chemical compounds, and natural product mixtures. Such agonists or antagonists so-identified may be natural or modified substrates, ligands, receptors, enzymes, etc., as the case may be, of the polypeptide; a structural or functional mimetic thereof (see Coligan *et al.*, Current Protocols in Immunology 1(2):Chapter 5 (1991)) or a small molecule. Such small molecules preferably have a molecular weight below 2,000 daltons, more preferably between 300 and 1,000 daltons, and most preferably between 400 and 700 daltons. It is preferred that these small molecules are organic molecules.

The screening method may simply measure the binding of a candidate compound to the polypeptide, or to cells or membranes bearing the polypeptide, or a fusion protein thereof, by means of a label directly or indirectly associated with the candidate compound. Alternatively, the screening method may involve measuring or detecting (qualitatively or quantitatively) the competitive binding of a candidate compound to the polypeptide against a labeled competitor (*e.g.* agonist or antagonist). Further, these screening methods may test whether the candidate compound results in a signal generated by activation or inhibition of the polypeptide, using detection systems appropriate to the cells bearing the polypeptide. Inhibitors of activation are generally assayed in the presence of a known agonist and the effect on activation by the agonist by the presence of the candidate compound is observed. Further, the screening methods may simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide of the present invention, to form a mixture, measuring a VANILREP6 activity in the mixture, and comparing the VANILREP6 activity of the mixture to a control mixture which contains no candidate compound.

Polypeptides of the present invention may be employed in conventional low capacity screening methods and also in high-throughput screening (HTS) formats. Such HTS formats include not only the well-established use of 96- and, more recently, 384-well micotiter plates but also emerging methods such as the nanowell method described by Schullek *et al.*, Anal Biochem., 246, 20-29, (1997).

Fusion proteins, such as those made from Fc portion and VANILREP6 polypeptide, as hereinbefore described, can also be used for high-throughput screening assays to identify antagonists for the polypeptide of the present invention (see D. Bennett *et al.*, J Mol Recognition, 8:52-58 (1995); and K. Johanson *et al.*, J Biol Chem, 270(16):9459-9471 (1995)).



The polynucleotides, polypeptides and antibodies to the polypeptide of the present invention may also be used to configure screening methods for detecting the effect of added compounds on the production of mRNA and polypeptide in cells. For example, an ELISA assay may be constructed for measuring secreted or cell associated levels of polypeptide using  
5 monoclonal and polyclonal antibodies by standard methods known in the art. This can be used to discover agents that may inhibit or enhance the production of polypeptide (also called antagonist or agonist, respectively) from suitably manipulated cells or tissues.

A polypeptide of the present invention may be used to identify membrane bound or soluble receptors, if any, through standard receptor binding techniques known in the art. These  
10 include, but are not limited to, ligand binding and crosslinking assays in which the polypeptide is labeled with a radioactive isotope (for instance,  $^{125}\text{I}$ ), chemically modified (for instance, biotinylated), or fused to a peptide sequence suitable for detection or purification, and incubated with a source of the putative receptor (cells, cell membranes, cell supernatants, tissue extracts, bodily fluids). Other methods include biophysical techniques such as surface plasmon resonance  
15 and spectroscopy. These screening methods may also be used to identify agonists and antagonists of the polypeptide that compete with the binding of the polypeptide to its receptors, if any. Standard methods for conducting such assays are well understood in the art.

Examples of antagonists of polypeptides of the present invention include antibodies or, in some cases, oligonucleotides or proteins that are closely related to the ligands, substrates, receptors,  
20 enzymes, etc., as the case may be, of the polypeptide, *e.g.*, a fragment of the ligands, substrates, receptors, enzymes, etc.; or a small molecule that bind to the polypeptide of the present invention but do not elicit a response, so that the activity of the polypeptide is prevented.

Screening methods may also involve the use of transgenic technology and VANILREP6 gene. The art of constructing transgenic animals is well established. For example, the  
25 VANILREP6 gene may be introduced through microinjection into the male pronucleus of fertilized oocytes, retroviral transfer into pre- or post-implantation embryos, or injection of genetically modified, such as by electroporation, embryonic stem cells into host blastocysts. Particularly useful transgenic animals are so-called "knock-in" animals in which an animal gene is replaced by the human equivalent within the genome of that animal. Knock-in transgenic animals  
30 are useful in the drug discovery process, for target validation, where the compound is specific for the human target. Other useful transgenic animals are so-called "knock-out" animals in which the expression of the animal ortholog of a polypeptide of the present invention and encoded by an endogenous DNA sequence in a cell is partially or completely annulled. The gene knock-out may be targeted to specific cells or tissues, may occur only in certain cells or tissues as a consequence  
35 of the limitations of the technology, or may occur in all, or substantially all, cells in the animal.

Transgenic animal technology also offers a whole animal expression-cloning system in which introduced genes are expressed to give large amounts of polypeptides of the present invention

Screening kits for use in the above described methods form a further aspect of the present invention. Such screening kits comprise:

- 5 (a) a polypeptide of the present invention;
  - (b) a recombinant cell expressing a polypeptide of the present invention;
  - (c) a cell membrane expressing a polypeptide of the present invention; or
  - (d) an antibody to a polypeptide of the present invention;
- which polypeptide is preferably that of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID
- 10 NO:10.

It will be appreciated that in any such kit, (a), (b), (c) or (d) may comprise a substantial component.

### **Glossary**

15 The following definitions are provided to facilitate understanding of certain terms used frequently hereinbefore.

“Antibodies” as used herein includes polyclonal and monoclonal antibodies, chimeric, single chain, and humanized antibodies, as well as Fab fragments, including the products of an Fab or other immunoglobulin expression library.

20 “Isolated” means altered “by the hand of man” from its natural state, *i.e.*, if it occurs in nature, it has been changed or removed from its original environment, or both. For example, a polynucleotide or a polypeptide naturally present in a living organism is not “isolated,” but the same polynucleotide or polypeptide separated from the coexisting materials of its natural state is “isolated”, as the term is employed herein. Moreover, a polynucleotide or polypeptide that is

25 introduced into an organism by transformation, genetic manipulation or by any other recombinant method is “isolated” even if it is still present in said organism, which organism may be living or non-living.

“Polynucleotide” generally refers to any polyribonucleotide (RNA) or polydeoxribonucleotide (DNA), which may be unmodified or modified RNA or DNA.

30 “Polynucleotides” include, without limitation, single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, “polynucleotide” refers to triple-stranded regions

35 comprising RNA or DNA or both RNA and DNA. The term “polynucleotide” also includes DNAs or RNAs containing one or more modified bases and DNAs or RNAs with backbones

modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications may be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically or metabolically modified forms of polynucleotides as typically found in nature, as well as the chemical forms of DNA and RNA characteristic of viruses and cells. "Polynucleotide" also embraces relatively short polynucleotides, often referred to as oligonucleotides.

"Polypeptide" refers to any polypeptide comprising two or more amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres. "Polypeptide" refers to both short chains, commonly referred to as peptides, oligopeptides or oligomers, and to longer chains, generally referred to as proteins. Polypeptides may contain amino acids other than the 20 gene-encoded amino acids. "Polypeptides" include amino acid sequences modified either by natural processes, such as post-translational processing, or by chemical modification techniques that are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications may occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present to the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched and branched cyclic polypeptides may result from post-translation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, biotinylation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphatidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cystine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination (see, for instance, *Proteins - Structure and Molecular Properties*, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York, 1993; Wold, F., *Post-translational Protein Modifications: Perspectives and Prospects*, 1-12, in *Post-translational Covalent Modification of Proteins*, B. C. Johnson, Ed., Academic Press, New York, 1983; Seifter *et al.*, "Analysis for protein modifications and nonprotein cofactors", *Meth Enzymol*, 182, 626-646, 1990, and Rattan *et al.*, "Protein Synthesis: Post-translational Modifications and Aging", *Ann NY Acad Sci*, 663, 48-62, 1992).

"Fragment" of a polypeptide sequence refers to a polypeptide sequence that is shorter than the reference sequence but that retains essentially the same biological function or activity as the reference polypeptide. "Fragment" of a polynucleotide sequence refers to a polynucleotide sequence that is shorter than the reference sequence of SEQ ID NO:1.

5       "Variant" refers to a polynucleotide or polypeptide that differs from a reference polynucleotide or polypeptide, but retains the essential properties thereof. A typical variant of a polynucleotide differs in nucleotide sequence from the reference polynucleotide. Changes in the nucleotide sequence of the variant may or may not alter the amino acid sequence of a polypeptide encoded by the reference polynucleotide. Nucleotide changes may result in amino acid  
10 substitutions, additions, deletions, fusions and truncations in the polypeptide encoded by the reference sequence, as discussed below. A typical variant of a polypeptide differs in amino acid sequence from the reference polypeptide. Generally, alterations are limited so that the sequences of the reference polypeptide and the variant are closely similar overall and, in many regions, identical. A variant and reference polypeptide may differ in amino acid sequence by one or more  
15 substitutions, insertions, deletions in any combination. A substituted or inserted amino acid residue may or may not be one encoded by the genetic code. Typical conservative substitutions include Gly, Ala; Val, Ile, Leu; Asp, Glu; Asn, Gln; Ser, Thr; Lys, Arg; and Phe and Tyr. A variant of a polynucleotide or polypeptide may be naturally occurring such as an allele, or it may be a variant that is not known to occur naturally. Non-naturally occurring variants of polynucleotides  
20 and polypeptides may be made by mutagenesis techniques or by direct synthesis. Also included as variants are polypeptides having one or more post-translational modifications, for instance glycosylation, phosphorylation, methylation, ADP ribosylation and the like. Embodiments include methylation of the N-terminal amino acid, phosphorylations of serines and threonines and modification of C-terminal glycines.

25       "Allele" refers to one of two or more alternative forms of a gene occurring at a given locus in the genome.

      "Polymorphism" refers to a variation in nucleotide sequence (and encoded polypeptide sequence, if relevant) at a given position in the genome within a population.

      "Single Nucleotide Polymorphism" (SNP) refers to the occurrence of nucleotide  
30 variability at a single nucleotide position in the genome, within a population. An SNP may occur within a gene or within intergenic regions of the genome. SNPs can be assayed using Allele Specific Amplification (ASA). For the process at least 3 primers are required. A common primer is used in reverse complement to the polymorphism being assayed. This common primer can be between 50 and 1500 bps from the polymorphic base. The other two (or more) primers are  
35 identical to each other except that the final 3' base wobbles to match one of the two (or more)

alleles that make up the polymorphism. Two (or more) PCR reactions are then conducted on sample DNA, each using the common primer and one of the Allele Specific Primers.

"Splice Variant" as used herein refers to cDNA molecules produced from RNA molecules initially transcribed from the same genomic DNA sequence but which have undergone alternative RNA splicing. Alternative RNA splicing occurs when a primary RNA transcript undergoes splicing, generally for the removal of introns, which results in the production of more than one mRNA molecule each of that may encode different amino acid sequences. The term splice variant also refers to the proteins encoded by the above cDNA molecules.

"Identity" reflects a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, determined by comparing the sequences. In general, identity refers to an exact nucleotide to nucleotide or amino acid to amino acid correspondence of the two polynucleotide or two polypeptide sequences, respectively, over the length of the sequences being compared.

"% Identity" - For sequences where there is not an exact correspondence, a "% identity" may be determined. In general, the two sequences to be compared are aligned to give a maximum correlation between the sequences. This may include inserting "gaps" in either one or both sequences, to enhance the degree of alignment. A % identity may be determined over the whole length of each of the sequences being compared (so-called global alignment), that is particularly suitable for sequences of the same or very similar length, or over shorter, defined lengths (so-called local alignment), that is more suitable for sequences of unequal length.

"Similarity" is a further, more sophisticated measure of the relationship between two polypeptide sequences. In general, "similarity" means a comparison between the amino acids of two polypeptide chains, on a residue by residue basis, taking into account not only exact correspondences between a between pairs of residues, one from each of the sequences being compared (as for identity) but also, where there is not an exact correspondence, whether, on an evolutionary basis, one residue is a likely substitute for the other. This likelihood has an associated "score" from which the "% similarity" of the two sequences can then be determined.

Methods for comparing the identity and similarity of two or more sequences are well known in the art. Thus for instance, programs available in the Wisconsin Sequence Analysis Package, version 9.1 (Devereux J et al, Nucleic Acids Res, 12, 387-395, 1984, available from Genetics Computer Group, Madison, Wisconsin, USA), for example the programs BESTFIT and GAP, may be used to determine the % identity between two polynucleotides and the % identity and the % similarity between two polypeptide sequences. BESTFIT uses the "local homology" algorithm of Smith and Waterman (J Mol Biol, 147,195-197, 1981, Advances in Applied Mathematics, 2, 482-489, 1981) and finds the best single region of similarity between two sequences. BESTFIT is more suited to comparing two polynucleotide or two polypeptide

sequences that are dissimilar in length, the program assuming that the shorter sequence represents a portion of the longer. In comparison, GAP aligns two sequences, finding a "maximum similarity", according to the algorithm of Needleman and Wunsch (J Mol Biol, 48, 443-453, 1970). GAP is more suited to comparing sequences that are approximately the same length and an alignment is expected over the entire length. Preferably, the parameters "Gap Weight" and "Length Weight" used in each program are 50 and 3, for polynucleotide sequences and 12 and 4 for polypeptide sequences, respectively. Preferably, % identities and similarities are determined when the two sequences being compared are optimally aligned.

Other programs for determining identity and/or similarity between sequences are also known in the art, for instance the BLAST family of programs (Altschul S F et al, J Mol Biol, 215, 403-410, 1990, Altschul S F et al, Nucleic Acids Res., 25:389-3402, 1997, available from the National Center for Biotechnology Information (NCBI), Bethesda, Maryland, USA and accessible through the home page of the NCBI at [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov)) and FASTA (Pearson W R, Methods in Enzymology, 183, 63-99, 1990; Pearson W R and Lipman D J, Proc Nat Acad Sci USA, 85, 2444-2448, 1988, available as part of the Wisconsin Sequence Analysis Package).

Preferably, the BLOSUM62 amino acid substitution matrix (Henikoff S and Henikoff J G, Proc. Nat. Acad Sci. USA, 89, 10915-10919, 1992) is used in polypeptide sequence comparisons including where nucleotide sequences are first translated into amino acid sequences before comparison.

Preferably, the program BESTFIT is used to determine the % identity of a query polynucleotide or a polypeptide sequence with respect to a reference polynucleotide or a polypeptide sequence, the query and the reference sequence being optimally aligned and the parameters of the program set at the default value, as hereinbefore described.

"Identity Index" is a measure of sequence relatedness which may be used to compare a candidate sequence (polynucleotide or polypeptide) and a reference sequence. Thus, for instance, a candidate polynucleotide sequence having, for example, an Identity Index of 0.95 compared to a reference polynucleotide sequence is identical to the reference sequence except that the candidate polynucleotide sequence may include on average up to five differences per each 100 nucleotides of the reference sequence. Such differences are selected from the group consisting of at least one nucleotide deletion, substitution, including transition and transversion, or insertion. These differences may occur at the 5' or 3' terminal positions of the reference polynucleotide sequence or anywhere between these terminal positions, interspersed either individually among the nucleotides in the reference sequence or in one or more contiguous groups within the reference sequence. In other words, to obtain a polynucleotide sequence having an Identity Index of 0.95 compared to a reference polynucleotide sequence, an average of up to 5 in every 100 of the nucleotides of the in the reference sequence may be deleted, substituted or inserted, or any

combination thereof, as hereinbefore described. The same applies *mutatis mutandis* for other values of the Identity Index, for instance 0.96, 0.97, 0.98 and 0.99.

Similarly, for a polypeptide, a candidate polypeptide sequence having, for example, an Identity Index of 0.95 compared to a reference polypeptide sequence is identical to the reference sequence except that the polypeptide sequence may include an average of up to five differences per each 100 amino acids of the reference sequence. Such differences are selected from the group consisting of at least one amino acid deletion, substitution, including conservative and non-conservative substitution, or insertion. These differences may occur at the amino- or carboxy-terminal positions of the reference polypeptide sequence or anywhere between these terminal positions, interspersed either individually among the amino acids in the reference sequence or in one or more contiguous groups within the reference sequence. In other words, to obtain a polypeptide sequence having an Identity Index of 0.95 compared to a reference polypeptide sequence, an average of up to 5 in every 100 of the amino acids in the reference sequence may be deleted, substituted or inserted, or any combination thereof, as hereinbefore described. The same applies *mutatis mutandis* for other values of the Identity Index, for instance 0.96, 0.97, 0.98 and 0.99.

The relationship between the number of nucleotide or amino acid differences and the Identity Index may be expressed in the following equation:

$$n_a \leq x_a - (x_a \bullet I),$$

in which:

$n_a$  is the number of nucleotide or amino acid differences,

$x_a$  is the total number of nucleotides or amino acids in SEQ ID NO:1 or SEQ ID NO:2, respectively,

$I$  is the Identity Index,

$\bullet$  is the symbol for the multiplication operator, and

in which any non-integer product of  $x_a$  and  $I$  is rounded down to the nearest integer prior to subtracting it from  $x_a$ .

"Homolog" is a generic term used in the art to indicate a polynucleotide or polypeptide sequence possessing a high degree of sequence relatedness to a reference sequence. Such relatedness may be quantified by determining the degree of identity and/or similarity between the two sequences as hereinbefore defined. Falling within this generic term are the terms "ortholog", and "paralog". "Ortholog" refers to a polynucleotide or polypeptide that is the functional equivalent of the polynucleotide or polypeptide in another species. "Paralog" refers to a polynucleotide or polypeptide that within the same species which is functionally similar.

"Fusion protein" refers to a protein encoded by two, often unrelated, fused genes or fragments thereof. In one example, EP-A-0 464 533-A discloses fusion proteins comprising

various portions of constant region of immunoglobulin molecules together with another human protein or part thereof. In many cases, employing an immunoglobulin Fc region as a part of a fusion protein is advantageous for use in therapy and diagnosis resulting in, for example, improved pharmacokinetic properties [see, *e.g.*, EP-A 0232 262]. On the other hand, for some uses it would be desirable to be able to delete the Fc part after the fusion protein has been expressed, detected and purified.

All publications and references, including but not limited to patents and patent applications, cited in this specification are herein incorporated by reference in their entirety as if each individual publication or reference were specifically and individually indicated to be incorporated by reference herein as being fully set forth. Any patent application to which this application claims priority is also incorporated by reference herein in its entirety in the manner described above for publications and references.

## Examples

Example 1. VANILREP6 is expressed in the nervous system.

Tissue and cell expression of human VANILREP6 was studied using TaqMan quantitative RT-PCR (Gibson et al., 1996) according to the manufacturers instructions. TaqMan reactions were conducted using probes for human GAPDH, cyclophilin and human VANILREP6. The human VANILREP6 probe consisted of 5'- CCTCCTCAACATGCTCATTGCT (SEQ ID NO:11) and 5'- ATGCGTTTCGCTCTCCTTGG (SEQ ID NO:12) flanking primers and a 5'- CGTTCTCCACAGTCTCGCCCATCA (SEQ ID NO:13) fluorogenic probe. Data were analysed using the Power Macintosh software accompanying the ABI Prism™ 7700.

Result: The data from a screen of body tissues shows that human VANILREP6 is most prominently expressed in nervous tissue. Analysis of brain regions shows uniform expression across a wide range of brain regions including spinal cord, cortex, hippocampus, thalamus, hypothalamus, amygdala, caudate nucleus and putamen. Expression in dorsal root ganglia was found to be three times that found in spinal cord or whole brain.

A screen of primary and clonal cell cultures shows significant expression in muscle cell lines, megakaryocyte cell lines, liver and kidney cell lines.

Table of relative mRNA expression, on a qualitative score from 1 to highest found 5.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
T	5	2	1	0	0	0	0	1	1	1	0	1	0	1	0	0	1	0	0	0
C	0	0	1	4	0	0	0	0	0	5	0	4	0	0	1	0	1	1	0	0

Where T the category of different body tissues, and

A CNS,

B pituitary,

C heart,

D lung,

E liver,

F foetal liver,

G kidney,

H skeletal

muscle,

5 I stomach,

J intestine,

K spleen,

L lymphocytes,

M macrophages,

N adipose,

O pancreas,

P prostate,

Q placenta,

R cartilage,

S bone,

T bone

marrow.

10 Where C the category of different cell lines, and

A aortic smooth muscle cells, B bladder smooth muscle cells, C C20A4,

D HOS,

E SAOS2,

F lymphocyte,

G macrophage,

H platelets,

I

neutrophil,

15 J M-07e,

K HepG2,

L HK-2,

M SK-N-MC,

N SK-N-SH,

O NT-2,

P 1321N1,

Q WRL68,

Rprimary

human chondrocytes,

S Hs-683,

T HEK293.

20

Levels of mRNA expression were studied across brain regions and in peripheral nervous tissue.

Table of relative mRNA expression, on a qualitative score from 1 to highest found 5

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
2	2	1	5	2	2	2	2	2	1	2	1	2	1	2	4

25 Where:

A amygdala

B caudate nucleus

C cerebellum

D corpus callosum

E temporal cortex

F hippocampus

G hypothalamus

H nuc. accumbens

I putamen

J sub. nigra

K thalamus

L fetal brain

30 M spinal cord

N pit. gland

O whole brain

**Example 2 – Activation of VANILREP6 by heat.**

5 Whole cell patch clamp recordings were performed essentially as described (Gunthorpe et al. 2000). Experiments were conducted at room temperature (20-24 °C) unless otherwise stated, Cells were plated onto glass coverslips coated with poly-D-lysine at a density of ~26,000 cells.cm<sup>-2</sup> and used after 1-3 days. The extracellular solution consisted of (mM) NaCl, 130; KCl, 5; CaCl<sub>2</sub>, 2; MgCl<sub>2</sub>, 1; Glucose, 30; HEPES-NaOH, 25; pH 7.3. Patch pipettes  
10 (resistance 2-5 MΩ) were fabricated on a Sutter instruments P-87 electrode puller and were filled with the following solution (mM): CsCl, 140; MgCl<sub>2</sub>, 4; EGTA, 10; HEPES-CsOH, 10; pH 7.3.

Drugs or solutions were applied using an automated device for fast switching of solutions (Warner Instruments SF-77B). To achieve rapid temperature jumps, the solution supplied to one  
15 of the bank of glass tubes in the solution exchanger was first passed through a solution heating device (Warner Instruments in-line heater SH-27A), allowing the temperature of the solution flowing through that barrel to be elevated in a controlled manner. The system was calibrated using two miniature thermocouples (~100 μM diameter, Omega Instruments) placed at the input to the glass tube through which the heated solution flowed and adjacent to its output in the  
20 location usually occupied by the recorded cell. Data acquisition and analysis were performed using the pClamp7 software suite and Origin (Microcal).

Whole-cell patch clamp recordings were made from HEK293 cells transiently transfected with human VR6 (co-transfected with a GFP vector to allow visualisation of transfected cells). Control recordings were made from cells similarly transfected with vector and GFP alone.  
25 Very small inward currents (<20 pA at 48 °C, n=4) were observed in control cells over the temperature range recorded (23-48 °C), which are likely due to changes in the physico-chemical properties of the cell membrane or the resistance of the seal between the electrode and the cell membrane (Cesare & McNaughton 1996; Hayes et al., 2000). In cells transfected with hVR6 an additional temperature-induced current component was observed in ~70% of cells studied (n=21).  
30 Temperature response curves established that this current exhibited a threshold for activation at ~39 °C and increased greatly in magnitude as the temperature was raised further. The VR6 heat-gated current was also associated with a large increase in current noise, indicative of the gating of a channel of relatively large single-channel conductance. No such effects were seen in the control cells studied.

Current-voltage relationships were established for heat-gated hVR6 using an appropriately timed voltage ramp protocol (-70 to +70 mV in 100 ms) applied during the response to heat (50-53 °C). The net hVR6 induced current was ascertained by subtraction of control data obtained during similar voltage ramps recorded at 23-25 °C. The hVR6 current-voltage relationship obtained  
5 exhibited a significant degree of outward rectification (rectification ration,  $I_{+70 \text{ mV}} / I_{-70 \text{ mV}}$  , of  $6.8 \pm 1.7$  fold, n=3), and a reversal potential close to 0 mV ( $-3.8 \pm 2.2$  mV, n=3), consistent with the gating of a non-selective cation channel and similar to the phenotype which is characteristic of the capsaicin receptor VR1.

## SEQUENCE INFORMATION

### SEQ ID NO:1

>

ATGGATTCCAACATCCGGCAGTGCATCTCTGGTAACTGTGATGACATGGACTCCCCCAG  
5 TCTCCTCARGATGATGTGACAGAGACCCCATCCAATCCCAACAGCCCCAGTGACACAGCTG  
GCCAAGGAAGAGCAGAGGAGGAAAAAGRGGCGGCTGAAGAAGCGCATCTTTGCAGCCGTG  
TCTGAGGGGCTGCGTGGAGGAGTTGGTAGAGTTGCTGGTGGAGCTGCAGGAGCTTTGCAGG  
CGGCGCCATGATGAGGATGTGCCTGACTTCCTCATGCACAAGCTGACGGCCTCCGACACG  
GGGAAGACCTGCCTGATGAAGGCCTTGTTAAACATCAACCCCCAACCAAGGAGATMGTG  
10 CGGATCCTGCTTGCCTTTGCTGAAGAGAACGACATCCTGGGCAGGTCATCAACGCCGAG  
TACACAGAGGAGGCTATGAAGGGCAGACGGGCGCTGAACATCGCCATCGAGCGGCGGCAG  
GGGGACATCGCAGCCCTGCTCATCGCCGCGGCGCGGACGTCAACGCGCACGCCAAGGGG  
GCCTTCTTCAACCCCCAAGTACCAACACGAAGGCTTCTACTTCGGTGAGACGCCCCCTGGCC  
CTGGCAGCATGCACCAACCAGCCCGAGATTGTGCAGCTGCTGATGGAGCACGAGCAGACG  
15 GACATCACCTCGCGGGACTCACGAGGCAACAACATCCTTCACGCCCTGGTGACCGTGGCC  
GAGGACTTCAAGACRCAGAATGACTTTGTGAAGCGCATGTACGACATGATCCTACTGCGG  
AGTGGCAACTGGGAGCTGGAGACCACTCGCAACAACGATGGCCTCACGCCGCTGCAGCTG  
GCCGCCAAGATGGGCAAGGCGGAGATCCTGAAGTACATCCTCAGTCGTGAGATCAAGGAG  
AAGCGGCTCCGGAGCCTGTCCAGGAAGTTCACCGACTGGGCGTACGGACCCGTGTCATCC  
20 TCCCTCTACGACCTCACCAACGTGGACACCACCACGGACAACCTCAGTGCTGGAAATCACT  
GTCTACAACACCAACATCGACAACCGGCATGAGATGCTGACCCTGGAGCCGCTGCACACG  
CTGCTGCATATGAAGTGGAAGAAGTTTGCCAAGCACATGTTCTTTCTGTCTTCTGCTTT  
TATTTCTTCTACAACATCACCTGACCCTCGTCTCGTACTACCGCCCCCGGGAGGAGGAG  
GCCATCCCGCACCCCTTGGCCCTGACGCACAAGATGGGGTGGCTGCAGCTCCTAGGGAGG  
25 ATGTTTGTGCTCATCTGGGCCATGTGCATCTCTGTGAAAGAGGGCATTGCCATCTTCCTG  
CTGAGACCCTCGGATCTGCAGTCCATCCTCTCGGATGCCTGGTTCCACTTTGTCTTTTTT  
ATCCAAGCTGTGCTTGTGATACTGTCTGTCTTCTTGTACTTGTTTGCCTACAAAGAGTAC  
CTCGCCTGCCTCGTGCTGGCCATGGCCCTGGGCTGGGCGAACATGCTCTACTATACGCGG  
GGTTTCCAGTCCATGGGCATGTACAGCGTCATGATCCAGAAGGTCATTTTGCATGATGTT  
30 CTGAAGTTCTTGTGTTGTATATATCGTGTTTTTGCTTGGATTGAGTAGCCTTGGCCTCG  
CTGATCGAGAAGTGTCCCAAAGACAACAAGGACTGCAGCTCCTACGGCAGCTTCAGCGAC  
GCAGTGCTGGAACCTTCAAGCTCACCATAGGCCTGGGTGACCTGAACATCCAGCAGAAC  
TCCAAGTATCCCATTCTCTTTCTGTTTCTGCTCATCACCTATGTCATCCTCACCTTTGTT  
CTCCTCCTCAACATGCTCATTGCTCTGATGGGCGAGACTGTGGAGAACGTCTCCAAGGAG  
35 AGCGAACGCATCTGGCGCCTGCAGAGAGCCAGGACCATCTTGGAGTTTGAGAAAATGTTA  
CCAGAATGGCTGAGGAGCAGATTCCGGATGGGAGAGCTGTGCAAAGTGGCCGAGGATGAT  
TTCCGACTGTGTTTGC GGATCAATGAGGTGAAGTGGACTGAATGGAAGACGCACGTCTCC  
TTCCTTAACGAAGACCCGGGGCCTGTAAGACGAACAGCAGATTTCAACAAAATCCAAGAT  
TCTTCCAGGAACAACAGCAAAACCACTCTCAATGCATTTGAAGAAGTCGAGGAATTCCCCG  
40 GAAACCTCGGTGTAG

### SEQ ID NO:2

>

MDSNIRQCISGNCDDMDSPQSPQDDVTETPSNPNSPSAQLAKEEQRRKKRRLKKRIFA  
SEGCVEELVELLVELQELCRRRHDEDVPDFLMHKLTASDTGKTCLMKALLNINPNTKEIV  
RILLAF AEENDILGRFINAEYTEEAYEGQTALNIAIERROGDIAALLIAAGADVNAHAKG  
5 AFFNPKYQHEGFYFGETPLALAACTNQPEIVQLLMEHEQTDITSRDSRGNNILHALVTVA  
EDFKTQND FVKRMYDMILLRSGNWELETTRNNDGLTPLQLAAKMGKAEILKYILSREIKE  
KRLRSLSRKFTDWAYGPVSSSLYDLTNVDTTDNSVLEITVYNTNIDNRHEMLTLEPLHT  
LLHMKWKKFAKHMFFLSFCFYFFYNITLTLVSYRPREEEAIPHPLALTHKMGWLQLLGR  
MFVLIWAMCISVKEGIAIFLLRPSDLQSILSDAWFHFVFFIQAVLVILSVFLYLFAFKEY  
10 LACLVLAMALGWANMLYYTRGFQSMGMYSVMIQKVLHDVLFVYIVFLLGFGVALAS  
LIEKCPKDNKDCSSYGSFSDAVLELFKLTIGLGLDNIQONSKYPILFLLITYVILTFV  
LLLNM LIALMGETVENVSKESERIWRLQRRARTILEFEKMLPEWLRSRFRMGELCKVAEDD  
FRLCLRINEVKWTEWKTHVSFLNEDPGPVRRTADFNKIQDSSRNNSKTTLNAFEVEVEEFP  
ETSV\*

15

**SEQ ID NO:3**

ATGGATTCCAACATCCGGCAGTGCATCTCTGGTAACTGTGATGACATGGACTCCCCCCAG  
TCTCCTCARGATGATGTGACAGAGACCCCATCCAATCCCAACAGCCCCAGTGACAGCTG  
GCCAAGGAAGAGCAGAGGAGGAAAAAGRGGCGGCTGAAGAAGCGCATCTTTGCAGCCGTG  
20 TCTGAGGGCTGCGTGGAGGAGTTGGTAGAGTTGCTGGTGGAGCTGCAGGAGCTTTGCAGG  
CGGCGCCATGATGAGGATGTGCCTGACTTCCTCATGCACAAGCTGACGGCCTCCGACACG  
GGGAAGACCTGCCTGATGAAGGCCTTGTAAACATCAACCCCAACACCAAGGAGATMGTG  
CGGATCCTGCTTGCCTTTGCTGAAGAGAACGACATCCTGGGCAGGTTCAACGCCCGAG  
TACACAGAGGAGGCCTATGAAGGGCAGACGGCGCTGAACATCGCCATCGAGCGGCGGCAG  
25 GGGGACATCGCAGCCCTGCTCATCGCCGCCGGCGCCGACGTCAACGCGCACGCCAAGGGG  
GCCTTCTTCAACCCCAAGTACCAACACGAAGGCTTCTACTTCGGTGAGACGCCCTGGCC  
CTGGCAGCATGCACCAACCAGCCCGAGATTGTGCAGCTGCTGATGGAGCACGAGCAGACG  
GACATCACCTCGCGGGACTCACGAGGCAACAACATCCTTCACGCCCTGGTGACCGTGCC  
GAGGACTTCAAGACRCAGAATGACTTTGTGAAGCGCATGTACGACATGATCCTACTGCGG  
30 AGTGGCAACTGGGAGCTGGAGACCACTCGCAACAACGATGGCCTCACGCCGCTGCAGCTG  
GCCGCCAAGATGGGCAAGGCGGAGATCCTGAAGTACATCCTCAGTCGTGAGATCAAGGAG  
AAGCGGCTCCGGAGCCTGTCCAGGAAGTTCACCGACTGGGCGTACGGACCCGTGTCATCC  
TCCCTCTACGACCTCACCAACGTGGACACCACCACGGACAACCTCAGTGCTGGAAATCACT  
GTCTACAACACCAACATCGACAACCGGCATGAGATGCTGACCCTGGAGCCGCTGCACACG  
35 CTGCTGCATATGAAGTGGAAGAAGTTTGCCAAGCACATGTTCTTTCTGTCTTCTGCTTT  
TATTTCTTCTACAACATCACCTGACCCTCGTCTCGTACTACCGCCCCCGGGAGGAGGAG  
GCCATCCCGCACCCCTTGGCCCTGACGCACAAGATGGGGTGGCTGCAGCTCCTAGGGAGG  
ATGTTTGTGCTCATCTGGGCCATGTGCATCTCTGTGAAAGAGGGCATTGCCATCTTCCTG  
CTGAGACCCTCGGATCTGCAGTCCATCCTCTCGGATGCCTGGTTCCACTTTGTCTTAGTA  
40 CCTCGCCTGCCTCGTGCTGGCCATGGCCCTGGGCTGGGCGAACATGCTCTACTATACGCG  
GGGTTTCCAGTCCATGGGCATGTACAGCGTCATGATCCAGAAGGTCATTTTGCATGA

**SEQ ID NO:4**

MDSNIRQCISGNCDDMDSPQSPQDDVTETPSNPNSPSAQLAKEEQRRKKRRLKKRIFA

SEGCVEELVELLVELQELCRRRHDEDVPDFLMHKLTASDTGKTCLMKALLNINPNTKEIV  
 RILLAF AEENDILGRFINAEYTEEAYEGQTALNIAIERRQGDIAALLIAAGADVNAHAKG  
 AFFNP KYQHEGFYFGETPLALAACTNQPEIVQLLMEHEQTDITSRDSRGNNILHALVTVA  
 EDFKTQND FVKRMYDMILLRSGNWELETTRNNDGLTPLQLAAKMGKAEILKYILSREIKE  
 5 KRLRSLSRKFTDWAYGPVSSSLYDLTNVDTTT DNSVLEITVYNTNIDNRHEMLTLEPLHT  
 LLHMKWKKFAKHMFFLSFCFYFFYNITLTLVSYYRPREEEAIPHPLALTHKMGWLQLLGR  
 MFVLIWAMCISVKEGIAIFLLRPSDLQSI LSDAWFHFVLPRLPRAGHGPGGLGEHALLYA  
 GFPVHGHVQRHDPEGHFA\*

**10 SEQ ID NO:5**

ATGAAAGCCCACCCCAAGGAGATGGTGCCTCTCATGGGCAAGAGAGTTGCTGCCCCCAGT  
 GGGAACCCTGCCGTCCTGCCAGAGAAGAGGCCGCGGAGATCACCCCCACAAAGAAGAGT  
 GCACACTTCTTCCTGGAGATAGAAGGGTTTGAACCCAACCCACAGTTGCCAAGACCTCT  
 CCTCCTGTCTTCTCCAAGCCCATGGATTCCAACATCCGGCAGTGCATCTCTGGTAACTGT  
 15 GATGACATGGACTCCCCCAGTCTCCTCARGATGATGTGACAGAGACCCCATCCAATCCC  
 AACAGCCCCAGTGCACAGCTGGCCAAGGAAGAGCAGAGGAGGAAAAAGRGGCGGCTGAAG  
 AAGCGCATCTTTGCAGCCGTGTCTGAGGGCTGCGTGGAGGAGTTGGTAGAGTTGCTGGTG  
 GAGCTGCAGGAGCTTTGCAGGCGGCGCCATGATGAGGATGTGCCTGACTTCCTCATGCAC  
 AAGCTGACGGCCTCCGACACGGGGAAGACCTGCCTGATGAAGGCCTTGTTAAACATCAAC  
 20 CCCAACACCAAGGAGATMGTGCGGATCCTGCTTGCTTTGCTGAAGAGAACGACATCCTG  
 GGCAGGTTTCATCAACGCCGAGTACACAGAGGAGGCCTATGAAGGGCAGACGGCGCTGAAC  
 ATCGCCATCGAGCGGCGGCAGGGGGACATCGCAGCCCTGCTCATCGCCGCCGGCGCCGAC  
 GTCAACGCGCACGCCAAGGGGGCCTTCTTCAACCCAAGTACCAACACGAAGGCTTCTAC  
 TTCGGTGAGACGCCCCTGGCCCTGGCAGCATGCACCAACCAGCCCGAGATTGTGCAGCTG  
 25 CTGATGGAGCACGAGCAGACGGACATCACCTCGCGGGACTCACGAGGCAACAACATCCTT  
 CACGCCCTGGTGACCGTGGCCGAGGACTTCAAGACRCAGAATGACTTTGTGAAGCGCATG  
 TACGACATGATCCTACTGCGGAGTGGCAACTGGGAGCTGGAGACCACTCGCAACAACGAT  
 GGCCTCACGCCGCTGCAGCTGGCCGCCAAGATGGGCAAGGCGGAGATCCTGAAGTACATC  
 CTCAGTCGTGAGATCAAGGAGAAGCGGCTCCGGAGCCTGTCCAGGAAGTTCACCGACTGG  
 30 GCGTACGGACCCGTGTCATCCTCCCTCTACGACCTCACCAACGTGGACACCACCACGGAC  
 AACTCAGTGCTGGAATCACTGTCTACAACACCAACATCGACAACCGGCATGAGATGCTG  
 ACCCTGGAGCCGCTGCACACGCTGCTGCATATGAAGTGGAAGAAGTTTGCCAAGCACATG  
 TTCTTTCTGTCTTCTGCTTTTATTTCTTCTACAACATCACCCCTGACCCTCGTCTCGTAC  
 TACCGCCCCCGGGAGGAGGAGGCCATCCCGCACCCCTTGGCCCTGACGCACAAGATGGGG  
 35 TGGCTGCAGCTCCTAGGGAGGATGTTTGTGCTCATCTGGGCCATGTGCATCTCTGTGAAA  
 GAGGGCATTGCCATCTTCCTGCTGAGACCCTCGGATCTGCAGTCCATCCTCTCGGATGCC  
 TGGTTCCACTTTGTCTTTTTTATCCAAGCTGTGCTTGTGATACTGTCTGTCTTCTGTAC  
 TTGTTTGCCTACAAAGAGTACCTCGCCTGCCTCGTGCTGGCCATGGCCCTGGGCTGGGCG  
 AACATGCTCTACTATACGCGGGGTTTCCAGTCCATGGGCATGTACAGCGTCATGATCCAG  
 40 AAGGTATTTTTGCATGATGTTCTGAAGTTCTTGTGTTGTATATATCGTGTTTTGCTTGGA  
 TTTGGAGTAGCCTTGGCCTCGCTGATCGAGAAGTGTCCCAAAGACAACAAGGACTGCAGC  
 TCCTACGGCAGCTTCAGYGACGAGTGCTGGAACCTTCAAGCTCACCATAGGCCTGGGT  
 GAYCTGAACATCCAGCAGAACTCCAAGTATCCCATTTCTCTTTCTGTTCTGCTCATCACC

TATGTCATCCTCACCTTTGTTCTCCTCCTCAACATGCTCATTGCTCTGATGGGCGAGACT  
 GTGGAGAACGTCTCCAAGGAGAGCGAACGCATCTGGCGCCTGCAGAGAGCCAGGACCATC  
 TTGGAGTTTGAGAAAATGTTACCAGAATGGCTGAGGAGCAGATTCCGGATGGGAGAGCTG  
 TGCAAAGTGGCCGAGGATGATTTCCGACTGTGTTTGCGGATCAATGAGGTGAAGTGGACT  
 5 GAATGGAAGACGCACGTCTCCTTTCCTTAACGAAGACCCGGGGCCTGTAAGACGAACAGAT  
 TTCAACAAAATCCAAGATTCTTCCAGGAACAACAGCAAAACCACTCTCAATGCATTTGAA  
 GAAGTCGAGGAATTCCCGGAAACCTCGGTGTAG

**SEQ ID NO:6**

10 MKAHPKEMVPLMGKRVAAPSGNPAVLPEKRPAEITPTKKS AHFFLEIEGFEPNPTVAKTS  
 PPVFSKPMDSNIRQCISGNCDDMDSPQSPQDDVTETPSNPNSPSAQLAKEEQRRKKGRLK  
 KRIFAAVSEGCVEELVELLVELQELCRRRHDEDVPDFLMHKLTASDTGKTCMLKALLNIN  
 PNTKEIVRILLAF AEENDILGRFINAEYTEEAYEGQTALNIAIERRQGDIAALLIAAGAD  
 VNAHAKGAFFNP KYQHEGFYFGETPLALAACTNQPEIVQLLMEHEQTDITSRDSRGNNIL  
 15 HALVTV AEDFKTQND FVKRMYDMILLRSGNWELETTRNNDGLT PLQLAAKMGA EILKYI  
 LSREIKEKRLRLSLSRKFTDWAYGPVSSSLYDLTNVDTTT DNSVLEITVYNTNIDNRHEML  
 TLEPLHTLLHMKWK KFAKHMFFLSFCFYFFYNITLTLVSYYRPREEEAIPHPLALTHKMG  
 WLQLLGRMFVLIWAMCISVKEGIAIFLLRPSDLQSILSDAWFHFVFFIQAVLVILSVFLY  
 LFAYKEYLACLVLAMALGWANMLYYTRGFQSMGMYSVMIQKVILHDVLKFLFVYIVFLLG  
 20 FGVALASLIEKCPKDNKDCSSYGSFSDAVLELFKLTIGLGLDNIQQNSKYPILFLFLLIT  
 YVILTFVLLNMLIALMGETVENVSKESERIWRLQRARTILEFEKMLPEWLRSRFRMGEL  
 CKVAEDDFRLCLRINEVKWTEWKTHVSFLNEDPGPVRRTDFNKIQDSSRNNSKTTLNAFE  
 EVEEFPETSV\*

**SEQ ID NO:7**

25 TTTTAATCTTGCTAATTAATTCTTGGAATAATCAGGAACGAAACAGACAACCTTTAAGAAA  
 ATATTGTTCTTACTTAGACTATACTGAACTGCTATGTGCCGGTGAAGAGAAGTYTGTATG  
 CCAGAGCGGCCGCTGAATTCTAGAAGCCGTCCTGCCAGAGAAGAGGCCGGCGGAGATCAC  
 CCCACAAAGAAGAGTGCACACTTCTTCTCTGGAGATAGAAGGGTTTGAACCCAACCCAC  
 30 AGTTGCCAAGACCTCTCCTCCTGTCTTCTCCAAGCCC

**SEQ ID NO:8**

CAGGTGGCTCAGCCAGTTCTGCCTCTGACGCCTCATTCCAGCCATCCCTCTGCCTGCAAT  
 GAGAGCTTCCCGCCGCCTCAGCCACAGTCCCACCCGGGGCCTTGGGCCCCAGACATGCG  
 35 GTGATCTCAGGGCAAGGGTTGCCACGACCACCCAGAACCTCACCAGCCATGAAAGCCCAC  
 CCAAGGAGATGGTGCCTCTCATGGGCAAGAGAGTTGCTGCCCCCAGTGGGAACCCTGCC  
 GTCTGCCAGAGAAGAGGCCGGCGGAGATCACCCACAAAGAAGAGTGCACACTTCTTC  
 CTGGAGATAGAAGGGTTTGAACCCAACCCACAGTTGCCAAGACCTCTCCTCCTGTCTTC  
 TCCAAGCCCATGGATTCCAACATCCGGCAGTGCATCTCTGGTAACTGTGATGACATGGAC  
 40 TCCCCCAGTCTCCTCARGATGATGTGACAGAGACCCCATCCAATCCCAACAGCCCCAGT  
 GCACAGCTGGCCAAGGAAGAGCAGAGGAGGAAAAAGRGCGGCTGAAGAAGCGCATCTTT  
 GCAGCCGTGTCTGAGGGCTGCGTGGAGGAGTTGGTAGAGTTGCTGGTGGAGCTGCAGGAG  
 CTTTGCAGGCGGCGCCATGATGAGGATGTGCCTGACTTCCTCATGCACAAGCTGACGGCC

TCCGACACGGGGAAGACCTGCCTGATGAAGGCCTTGTTAAACATCAACCCCAACACCAAG  
 GAGATMGTGCGGATCCTGCTTGCCTTTGCTGAAGAGAACGACATCCTGGGCAGGTTTCATC  
 AACGCCGAGTACACAGAGGAGGCCTATGAAGGGCAGACGGCGCTGAACATCGCCATCGAG  
 CGGCGGCAGGGGGACATCGCAGCCCTGCTCATCGCCGCCGGCGCCGACGTCAACGCGCAC  
 5 GCCAAGGGGGCCTTCTTCAACCCCAAGTACCAACACGAAGGCTTCTACTTCGGTGAGACG  
 CCCCTGGCCCTGGCAGCATGCACCAACCAGCCCGAGATTGTGCAGCTGCTGATGGAGCAC  
 GAGCAGACGGACATCACCTCGCGGGACTCACGAGGCAACAACATCCTTCACGCCCTGGTG  
 ACCGTGGCCGAGGACTTCAAGACRCAGAATGACTTTGTGAAGCGCATGTACGACATGATC  
 CTACTGCGGAGTGGAACCTGGGAGCTGGAGACCACTCGCAACAACGATGGCCTCACGCCG  
 10 CTGCAGCTGGCCGCCAAGATGGGCAAGGCGGAGATCCTGAAGTACATCCTCAGTCGTGAG  
 ATCAAGGAGAAGCGGCTCCGGAGCCTGTCCAGGAAGTTCACCGACTGGGCGTACGGACCC  
 GTGTCATCCTCCCTCTACGACCTCACCAACGTGGACACCACCACGGACAACCTCAGTGCTG  
 GAAATCACTGTCTACAACACCAACATCGACAACCGGCATGAGATGCTGACCCTGGAGCCG  
 CTGCACACGCTGCTGCATATGAAGTGGAAGAAGTTTGCCAAGCACATGTTCTTTCTGTCC  
 15 TTCTGCTTTTATTTCTTCTACAACATCACCCCTGACCCTCGTCTCGTACTACCGCCCCCGG  
 GAGGAGGAGGCCATCCCGCACCCCTTGGCCCTGACGCACAAGATGGGGTGGCTGCAGCTC  
 CTAGGGAGGATGTTTGTGCTCATCTGGGCCATGTGCATCTCTGTGAAAGAGGGGCATTGCC  
 ATCTTCCTGCTGAGACCCTCGGATCTGCAGTCCATCCTCTCGGATGCCTGGTTCCACTTT  
 GTCTTTTTTATCCAAGCTGTGCTTGTGATACTGTCTGTCTTCTTGTACTTGTTTGCCTAC  
 20 AAAGAGTACCTCGCCTGCCTCGTGCTGGCCATGGCCCTGGGCTGGGCGAACATGCTCTAC  
 TATACGCGGGGTTTCCAGTCCATGGGCATGTACAGCGTCATGATCCAGAAGGTYATTTTG  
 CATGATGTTCTGAAGTTCTTGTTTGTATATATCGTGTTTTTGCTTGGATTTGGAGTAGCC  
 TTGGCCTCGCTGATCGAGAAGTGTCACCAAGACAACAAGGACTGCAGCTCCTACGGCAGC  
 TTCAGYGACGCAGTGCTGGAACCTCTTCAAGCTCACCATAGGCCTGGGTGAYCTGAACATC  
 25 CAGCAGAACTCCAAGTATCCCATTTCTTTCTGTTCTGCTCATCACCTATGTCATCCTC  
 ACCTTTGTTCTCCTCCTCAACATGCTCATTGCTCTGATGGGCGAGACTGTGGAGAACGTC  
 TCCAAGGAGAGCGAACGCATCTGGCGCCTGCAGAGAGCCAGGACCATCTTGGAGTTTGAG  
 AAAATGTTACCAGAATGGCTGAGGAGCAGATTCCGGATGGGAGAGCTGTGCAAAGTGGCC  
 GAGGATGATTTCCGACTGTGTTTTCGGATCAATGAGGTGAAGTGGACTGAATGGAAGACG  
 30 CACGTCTCCTTCTTAACGAAGACCCGGGGCCTGTAAGACGAACAGATTTCAACAAAATC  
 CAAGATTCTTCCAGGAACAACAGCAAAACCACTCTCAATGCATTTGAAGAAGTCGAGGAA  
 TTCCCGGAAACCTCGGTGTAGAAGCGGAACCCAGAGCTGGTGTGCGCGTGCGCTGTCTGG  
 CGCTGCAGGCGGAGTCACCGACTCTGTGCAGA

35 **SEQ ID NO:9**  
 ATGAAAGCCCACCCCAAGGAGATGGTGCCTCTCATGGGCAAGAGAGTTGCTGCCCCCAGT  
 GGGAACCCTGCCGTCTGCCAGAGAAGAGGCCGGCGGAGATCACCCCCACAAAGAAGAGT  
 GCACACTTCTTCTGAGATAGAAGGGTTTGAACCCAACCCACAGTTGCCAAGACCTCT  
 CCTCCTGTCTTCTCAAGCCCATGGATTCCAACATCCGGCAGTGTGCACAGCTGGCCAAG  
 40 GAAGAGCAGAGGAGGAAAAAGRGCGGCTGAAGAAGCGCATCTTTGCAGCCGTGTCTGAG  
 GGCTGCGTGGAGGAGTTGGTAGAGTTGCTGGTGGAGCTGCAGGAGCTTTGCAGGCGGCGC  
 CATGATGAGGATGTGCCTGACTTCTCATGCACAAGCTGACGGCCTCCGACACGGGGAAG  
 ACCTGCCTGATGAAGGCCTTGTTAAACATCAACCCCAACACCAAGGAGATMGTGCGGATC



CTGCTTGCCTTTGCTGAAGAGAACGACATCCTGGGCAGGTTTCATCAACGCCGAGTACACA  
 GAGGAGGCCTATGAAGGGCAGACGGCGCTGAACATCGCCATCGAGCGGCGGCAGGGGGAC  
 ATCGCAGCCCTGCTCATCGCCGCCGGCGCCGACGTCAACGCGCACGCCAAGGGGGCCTTC  
 TTCAACCCCAAGTACCAACACGAAGGCTTCTACTTCGGTGAGACGCCCCTGGCCCTGGCA  
 5 GCATGCACCAACCAGCCCGAGATTGTGCAGCTGCTGATGGAGCACGAGCAGACGGACATC  
 ACCTCGCGGGACTCACGAGGCAACAACATCCTTCACGCCCTGGTGACCGTGGCCGAGGAC  
 TTCAAGACRCAGAATGACTTTGTGAAGCGCATGTACGACATGATCCTACTGCGGAGTGGC  
 AACTGGGAGCTGGAGACCACTCGCAACAACGATGGCCTCACGCCGCTGCAGCTGGCCGCC  
 AAGATGGGCAAGGCGGAGATCCTGAAGTACATCCTCAGTCGTGAGATCAAGGAGAAGCGG  
 10 CTCCGGAGCCTGTCCAGGAAGTTCACCGACTGGGCGTACGGACCCGTGTCATCCTCCCTC  
 TACGACCTCACCAACGTGGACACCACCACGGACAACCTCAGTGCTGGAAATCACTGTCTAC  
 AACACCAACATCGACAACCGGCATGAGATGCTGACCCTGGAGCCGCTGCACACGCTGCTG  
 CATATGAAGTGAAGAAGTTTGCCAAGCACATGTTCTTTCTGTCTTCTGCTTTTATTTT  
 TTCTACAACATCACCTGACCCTCGTCTCGTACTACCGCCCCCGGGAGGAGGAGGCCATC  
 15 CCGCACCCCTTGGCCCTGACGCACAAGATGGGGTGGCTGCAGCTCCTAGGGAGGATGTTT  
 GTGCTCATCTGGGCCATGTGCATCTCTGTGAAAGAGGGCATTGCCATCTTCCTGCTGAGA  
 CCCTCGGATCTGCAGTCCATCCTCTCGGATGCCTGGTTCCACTTTGTCTTTTTTATCCAA  
 GCTGTGCTTGTGATACTGTCTGTCTTCTTGTACTTGTTCCTACAAAGAGTACCTCGCC  
 TGCCTCGTGCTGGCCATGGCCCTGGGCTGGGCGAACATGCTCTACTATACGCGGGGTTTC  
 20 CAGTCCATGGGCATGTACAGCGTCATGATCCAGAAGGTATTTTGCATGATGTTCTGAAG  
 TTCTTGTTTGTATATATCGTGTTTTTGGCTTGGATTTGGAGTAGCCTTGGCCTCGCTGATC  
 GAGAAGTGTCCCAAAGACAACAAGGACTGCAGCTCCTACGGCAGCTTCAGYGACGCAGTG  
 CTGGAACCTCTTCAAGCTCACCATAGGCCTGGGTGAYCTGAACATCCAGCAGAACTCCAAG  
 TATCCCATTTCTCTTTCTGTTCCTGCTCATCACCTATGTCATCCTCACCTTTGTTCTCCTC  
 25 CTCAACATGCTCATTTGCTCTGATGGGCGAGACTGTGGAGAACGTCTCCAAGGAGAGCGAA  
 CGCATCTGGCGCCTGCAGAGAGCCAGGACCATCTTGGAGTTTGAGAAAATGTTACCAGAA  
 TGGCTGAGGAGCAGATTCCGGATGGGAGAGCTGTGCAAAGTGGCCGAGGATGATTTCCGA  
 CTGTGTTTGGCGATCAATGAGGTGAAGTGGACTGAATGGAAGACGCACGTCTCCTTCCTT  
 AACGAAGACCCGGGGCCTGTAAGACGAACAGATTTCAACAAAATCCAAGATTCTTCCAGG  
 30 AACACAGCAAAACCACTCTCAATGCATTTGAAGAAGTCGAGGAATTCCCGGAAACCTCG  
 GTGTAG

**SEQ ID NO:10**

MKAHPKEMVPLMGKRVAAPSGNPAVLPEKRPAEITPTKKS AHFFLEIEGFEPNPTVAKTS  
 35 PPVFSKPMDSNIRQCAQLAKEEQRRKKGRLKKRIFA AVSEGCV EELVELLVELQELCRRR  
 HDEDVPDFLMHKLTASDTGKTCLMKALLNINPNTKEIVRILLAF AEENDILGRFINAEYT  
 EEAYEGQTALNIAIERRQGDIAALLIAAGADVNAHAKGAFFNP KYQHEGFYFGETPLALA  
 ACTNQPEIVQLLMEHEQTDITSRDSRGNNILHALVTVAEDFKT QNDFVKRMYDMILLRSG  
 NWELETTRNNDGLTPLQLAAKMGA EILKYILSREIKEKRLRSLSRKFTDWAYGPVSSSL  
 40 YDLTNVDTTT DNSVLEITVYNTNIDNRHEMLTLEPLHTLLHMKWKKFAKHMFFLSFCFYF  
 FYNITLTLVSYYRPREEEAI PHPLALTHKMGWLQLLGRMFVLIWAMCISVKEGIAIFLLR  
 PSDLQSI LSDAWFHVFVFIQAVLVILSVFLYLFAYKEYLACLVLAMALGWANMLYYTRGF  
 QSMGMYSVMIQKVILHDVLKFLFVYIVFLLGFGVALASLIEKCPKDNKDCSSYGSFSDAV

LELFKLTIGLGDLNIQQNSKYPILFLFLLITYVILTFVLLLNMLIALMGETVENVSKESE  
RIWRLQRARTILEFEKMLPEWLSRFRMGELCKVAEDDFRLCLRINEVKWTEWKTHVSFL  
NEDPGPVRRTDFNKIQDSSRNNSKTTLNAFEEVEEFPETSV\*

5

# SEQUENCE LISTING

5 <110> SmithKline Beecham plc  
SmithKline Beecham Corp.

<120> Novel Compounds

10 <130> GP30241

<160> 13

15 <170> FastSEQ for Windows Version 3.0

<210> 1

<211> 2175

<212> DNA

<213> Homo sapiens

20 <400> 1

atggattcca	acatccggca	gtgcatctct	ggtaactgtg	atgacatgga	ctccccccag	60
tctcctcarg	atgatgtgac	agagacccca	tccaatccca	acagccccag	tgacacagctg	120
gccaaggaag	agcagaggag	gaaaaagrpg	cggctgaaga	agcgcacctt	tgacagccgtg	180
tctgagggct	gcgtggagga	gttggtagag	ttgctggtgg	agctgcagga	gctttgcagg	240
cggcgccatg	atgaggatgt	gcctgacttc	ctcatgcaca	agctgacggc	ctccgacacg	300
gggaagacct	gcctgatgaa	ggccttggtt	aacatcaacc	ccaacaccaa	ggagatmggtg	360
cggatcctgc	ttgcctttgc	tgaagagaac	gacatcctgg	gcaggttcat	caacgccgag	420
tacacagagg	aggcctatga	agggcagacg	gcgctgaaca	tcgccatcga	gcggcggcag	480
ggggacatcg	cagccctgct	catcgccgcc	ggcgccgacg	tcaacgcgca	cgccaagggg	540
gccttcttca	accccaagta	ccaacacgaa	ggcttctact	tcggtgagac	gcccctggcc	600
ctggcagcat	gcaccaacca	gcccagagatt	gtgcagctgc	tgatggagca	cgagcagacg	660
gacatcacct	cgcgggactc	acgaggcaac	aacatccttc	acgccttgtt	gaccgtggcc	720
gaggacttca	agacrcagaa	tgactttgtg	aagcgcatgt	acgacatgat	cctactgcgg	780
agtggaact	gggagctgga	gaccactcgc	aacaacgatg	gcctcacgcc	gctgcagctg	840
gccgccaaga	tgggcaaggc	ggagatcctg	aagtacatcc	tcagtcgtga	gatcaaggag	900
aagcggctcc	ggagcctgtc	caggaagttc	accgactggg	cgtacggacc	cgtgtcatcc	960
tccctctacg	acctcaccaa	cgtggacacc	accacggaca	actcagtgtc	ggaaatcact	1020
gtctacaaca	ccaacatcga	caaccggcat	gagatgctga	ccctggagcc	gctgcacacg	1080
ctgctgcata	tgaagtggaa	gaagtttgcc	aagcacatgt	tctttctgtc	cttctgcttt	1140
tattttctct	acaacatcac	cctgaccctc	gtctcgtact	accgcccccg	ggaggaggag	1200
gcatcccgcc	accccttggc	cctgacgcac	aagatggggg	ggctgcagct	cctagggagg	1260
atgtttgtgc	tcctctgggc	catgtgcac	tctgtgaaag	agggcattgc	catcttcctg	1320
ctgagaccct	cggatctgca	gtccatcctc	tcggatgcct	ggttccactt	tgtctttttt	1380

atccaagctg tgcttgtgat actgtctgtc ttcttgtact tgtttgccta caaagagtac 1440  
 ctgcctgcc tcgtgctggc catggccctg ggctgggcca acatgctcta ctatacgcg 1500  
 ggtttccagt ccatgggcat gtacagcgtc atgatccaga aggtcatttt gcatgatgtt 1560  
 ctgaagttct tgtttgtata tatcgtgttt ttgcttggat ttggagtagc cttggcctcg 1620  
 5 ctgatcgaga agtgtcccaa agacaacaag gactgcagct cctacggcag cttcagcgac 1680  
 gcagtgcctg aactcttcaa gctcaccata ggctgggtg acctgaacat ccagcagaac 1740  
 tccaagtatc ccattctctt tctgttcttg ctcacacct atgtcatcct cacctttggt 1800  
 ctctctctca acatgctcat tgctctgatg ggcgagactg tggagaacgt ctccaaggag 1860  
 agcgaacgca tctggcgcct gcagagagcc aggaccatct tggagtgtga gaaaatgtta 1920  
 10 ccagaatggc tgaggagcag attccggatg ggagagctgt gcaaagtggc cgaggatgat 1980  
 ttccgactgt gtttgcggt caatgaggtg aagtggactg aatggaagac gcacgtctcc 2040  
 ttcttaacg aagaccggg gcctgtaaga cgaacagcag atttcaacaa aatccaagat 2100  
 tcttcagga acaacagcaa aaccactctc aatgcatttg aagaagtcga ggaattcccg 2160  
 gaaacctcgg tgtag 2175  
 15  
 <210> 2  
 <211> 724  
 <212> PRT  
 <213> Homo sapiens  
 20  
 <400> 2  
 Met Asp Ser Asn Ile Arg Gln Cys Ile Ser Gly Asn Cys Asp Asp Met  
 1 5 10 15  
 Asp Ser Pro Gln Ser Pro Gln Asp Asp Val Thr Glu Thr Pro Ser Asn  
 25 20 25 30  
 Pro Asn Ser Pro Ser Ala Gln Leu Ala Lys Glu Glu Gln Arg Arg Lys  
 35 40 45  
 Lys Arg Arg Leu Lys Lys Arg Ile Phe Ala Ala Val Ser Glu Gly Cys  
 50 55 60  
 30 Val Glu Glu Leu Val Glu Leu Leu Val Glu Leu Gln Glu Leu Cys Arg  
 65 70 75 80  
 Arg Arg His Asp Glu Asp Val Pro Asp Phe Leu Met His Lys Leu Thr  
 85 90 95  
 Ala Ser Asp Thr Gly Lys Thr Cys Leu Met Lys Ala Leu Leu Asn Ile  
 35 100 105 110  
 Asn Pro Asn Thr Lys Glu Ile Val Arg Ile Leu Leu Ala Phe Ala Glu  
 115 120 125  
 Glu Asn Asp Ile Leu Gly Arg Phe Ile Asn Ala Glu Tyr Thr Glu Glu  
 130 135 140  
 40 Ala Tyr Glu Gly Gln Thr Ala Leu Asn Ile Ala Ile Glu Arg Arg Gln  
 145 150 155 160  
 Gly Asp Ile Ala Ala Leu Leu Ile Ala Ala Gly Ala Asp Val Asn Ala  
 165 170 175  
 His Ala Lys Gly Ala Phe Phe Asn Pro Lys Tyr Gln His Glu Gly Phe

		180		185		190										
	Tyr	Phe	Gly	Glu	Thr	Pro	Leu	Ala	Leu	Ala	Ala	Cys	Thr	Asn	Gln	Pro
		195		200		205										
5	Glu	Ile	Val	Gln	Leu	Leu	Met	Glu	His	Glu	Gln	Thr	Asp	Ile	Thr	Ser
		210		215		220										
	Arg	Asp	Ser	Arg	Gly	Asn	Asn	Ile	Leu	His	Ala	Leu	Val	Thr	Val	Ala
	225			230		235										240
	Glu	Asp	Phe	Lys	Thr	Gln	Asn	Asp	Phe	Val	Lys	Arg	Met	Tyr	Asp	Met
				245		250										255
10	Ile	Leu	Leu	Arg	Ser	Gly	Asn	Trp	Glu	Leu	Glu	Thr	Thr	Arg	Asn	Asn
				260		265										270
	Asp	Gly	Leu	Thr	Pro	Leu	Gln	Leu	Ala	Ala	Lys	Met	Gly	Lys	Ala	Glu
				275		280										285
	Ile	Leu	Lys	Tyr	Ile	Leu	Ser	Arg	Glu	Ile	Lys	Glu	Lys	Arg	Leu	Arg
15				290		295										300
	Ser	Leu	Ser	Arg	Lys	Phe	Thr	Asp	Trp	Ala	Tyr	Gly	Pro	Val	Ser	Ser
	305					310						315				320
	Ser	Leu	Tyr	Asp	Leu	Thr	Asn	Val	Asp	Thr	Thr	Thr	Asp	Asn	Ser	Val
				325		330										335
20	Leu	Glu	Ile	Thr	Val	Tyr	Asn	Thr	Asn	Ile	Asp	Asn	Arg	His	Glu	Met
				340		345										350
	Leu	Thr	Leu	Glu	Pro	Leu	His	Thr	Leu	Leu	His	Met	Lys	Trp	Lys	Lys
				355		360										365
	Phe	Ala	Lys	His	Met	Phe	Phe	Leu	Ser	Phe	Cys	Phe	Tyr	Phe	Phe	Tyr
25				370		375										380
	Asn	Ile	Thr	Leu	Thr	Leu	Val	Ser	Tyr	Tyr	Arg	Pro	Arg	Glu	Glu	Glu
	385					390						395				400
	Ala	Ile	Pro	His	Pro	Leu	Ala	Leu	Thr	His	Lys	Met	Gly	Trp	Leu	Gln
				405		410										415
30	Leu	Leu	Gly	Arg	Met	Phe	Val	Leu	Ile	Trp	Ala	Met	Cys	Ile	Ser	Val
				420		425										430
	Lys	Glu	Gly	Ile	Ala	Ile	Phe	Leu	Leu	Arg	Pro	Ser	Asp	Leu	Gln	Ser
				435		440										445
	Ile	Leu	Ser	Asp	Ala	Trp	Phe	His	Phe	Val	Phe	Phe	Ile	Gln	Ala	Val
35				450		455										460
	Leu	Val	Ile	Leu	Ser	Val	Phe	Leu	Tyr	Leu	Phe	Ala	Tyr	Lys	Glu	Tyr
	465					470							475			480
	Leu	Ala	Cys	Leu	Val	Leu	Ala	Met	Ala	Leu	Gly	Trp	Ala	Asn	Met	Leu
				485		490										495
40	Tyr	Tyr	Thr	Arg	Gly	Phe	Gln	Ser	Met	Gly	Met	Tyr	Ser	Val	Met	Ile
				500		505										510
	Gln	Lys	Val	Ile	Leu	His	Asp	Val	Leu	Lys	Phe	Leu	Phe	Val	Tyr	Ile
				515		520										525
	Val	Phe	Leu	Leu	Gly	Phe	Gly	Val	Ala	Leu	Ala	Ser	Leu	Ile	Glu	Lys

	530		535		540
	Cys Pro Lys Asp Asn Lys Asp Cys Ser Ser Tyr Gly Ser Phe Ser Asp				
	545		550		555
	Ala Val Leu Glu Leu Phe Lys Leu Thr Ile Gly Leu Gly Asp Leu Asn				560
5		565		570	575
	Ile Gln Gln Asn Ser Lys Tyr Pro Ile Leu Phe Leu Phe Leu Leu Ile				
		580		585	590
	Thr Tyr Val Ile Leu Thr Phe Val Leu Leu Leu Asn Met Leu Ile Ala				
		595		600	605
10	Leu Met Gly Glu Thr Val Glu Asn Val Ser Lys Glu Ser Glu Arg Ile				
	610		615		620
	Trp Arg Leu Gln Arg Ala Arg Thr Ile Leu Glu Phe Glu Lys Met Leu				
	625		630		635
	Pro Glu Trp Leu Arg Ser Arg Phe Arg Met Gly Glu Leu Cys Lys Val				
15		645		650	655
	Ala Glu Asp Asp Phe Arg Leu Cys Leu Arg Ile Asn Glu Val Lys Trp				
		660		665	670
	Thr Glu Trp Lys Thr His Val Ser Phe Leu Asn Glu Asp Pro Gly Pro				
		675		680	685
20	Val Arg Arg Thr Ala Asp Phe Asn Lys Ile Gln Asp Ser Ser Arg Asn				
	690		695		700
	Asn Ser Lys Thr Thr Leu Asn Ala Phe Glu Glu Val Glu Glu Phe Pro				
	705		710		715
					720
	Glu Thr Ser Val				

25

	<210> 3
	<211> 1497
	<212> DNA
30	<213> Homo sapiens

	<400> 3	
	atggattcca acatccggca gtgcatctct ggtaactgtg atgacatgga ctccccccag	60
	tctcctcarg atgatgtgac agagacccca tccaatccca acagccccag tgcacagctg	120
35	gccaaggaag agcagaggag gaaaaagrpg cggtgaaga agcgcatctt tgcagccgtg	180
	tctgagggct gcgtggagga gttggtagag ttgctggtgg agctgcagga gctttgcagg	240
	cggcgccatg atgaggatgt gcctgacttc ctcatgcaca agctgacggc ctccgacacg	300
	gggaagacct gcctgatgaa ggccttgcta aacatcaacc ccaacaccaa ggagatmggtg	360
	cggatcctgc ttgcctttgc tgaagagaac gacatcctgg gcaggttcat caacgccgag	420
40	tacacagagg aggcctatga agggcagacg gcgtgaaca tcgccatcga gcggcggcag	480
	ggggacatcg cagccctgct catgcgcgcc ggcgccgacg tcaacgcgca cgccaagggg	540
	gcctttcttca accccaagta ccaacacgaa ggcttctact tcggtgagac gcccttgccc	600
	ctggcagcat gcaccaacca gcccgagatt gtgcagctgc tgatggagca cgagcagacg	660
	gacatcacct cgcgggactc acgaggcaac aacatccttc acgccttggt gaccgtggcc	720

5 gaggacttca agacrcagaa tgactttgtg aagcgcatgt acgacatgat cctactgogg 780  
 agtggcaact gggagctgga gaccactcgc aacaacgatg gcctcacgcc gctgcagctg 840  
 gccgccaaga tgggcaaggc ggagatcctg aagtacalcc tcagtcgtga gatcaaggag 900  
 aagcggctcc ggagcctgtc caggaagttc accgactggg cgtacggacc cgtgtcatcc 960  
 tccctctacg acctcaccaa cgtggacacc accacggaca actcagtgtt ggaaatcact 1020  
 gtctacaaca ccaacatcga caaccggcat gagatgctga ccctggagcc gctgcacacg 1080  
 ctgctgcata tgaagtggaa gaagtttgcc aagcacatgt tctttctgtc cttctgcttt 1140  
 tattttctct acaacatcac cctgaccctc gtctcgtact accgcccccg ggaggaggag 1200  
 gccatccccg accccttggc cctgacgcac aagatggggg ggctgcagct cctagggagg 1260  
 10 atgtttgtgc tcactctggc catgtgcac tctgtgaaag agggcattgc catcttctctg 1320  
 ctgagaccct cggatctgca gtccatcctc tcggatgcct ggttccactt tgtcttagta 1380  
 cctgcctgc ctcgtgctgg ccatggccct gggctgggcg aacatgctct actatacgcg 1440  
 gggttccag tccatgggca tgtacagcgt catgatccag aaggtcattt tgcata 1497

15 <210> 4  
 <211> 498  
 <212> PRT  
 <213> Homo sapiens

20 <400> 4  
 Met Asp Ser Asn Ile Arg Gln Cys Ile Ser Gly Asn Cys Asp Asp Met  
 1 5 10 15  
 Asp Ser Pro Gln Ser Pro Gln Asp Asp Val Thr Glu Thr Pro Ser Asn  
 20 25 30  
 25 Pro Asn Ser Pro Ser Ala Gln Leu Ala Lys Glu Glu Gln Arg Arg Lys  
 35 40 45  
 Lys Arg Arg Leu Lys Lys Arg Ile Phe Ala Ala Val Ser Glu Gly Cys  
 50 55 60  
 Val Glu Glu Leu Val Glu Leu Leu Val Glu Leu Gln Glu Leu Cys Arg  
 65 70 75 80  
 30 Arg Arg His Asp Glu Asp Val Pro Asp Phe Leu Met His Lys Leu Thr  
 85 90 95  
 Ala Ser Asp Thr Gly Lys Thr Cys Leu Met Lys Ala Leu Leu Asn Ile  
 100 105 110  
 35 Asn Pro Asn Thr Lys Glu Ile Val Arg Ile Leu Leu Ala Phe Ala Glu  
 115 120 125  
 Glu Asn Asp Ile Leu Gly Arg Phe Ile Asn Ala Glu Tyr Thr Glu Glu  
 130 135 140  
 Ala Tyr Glu Gly Gln Thr Ala Leu Asn Ile Ala Ile Glu Arg Arg Gln  
 145 150 155 160  
 40 Gly Asp Ile Ala Ala Leu Leu Ile Ala Ala Gly Ala Asp Val Asn Ala  
 165 170 175  
 His Ala Lys Gly Ala Phe Phe Asn Pro Lys Tyr Gln His Glu Gly Phe  
 180 185 190

	Tyr	Phe	Gly	Glu	Thr	Pro	Leu	Ala	Leu	Ala	Ala	Cys	Thr	Asn	Gln	Pro
			195					200					205			
	Glu	Ile	Val	Gln	Leu	Leu	Met	Glu	His	Glu	Gln	Thr	Asp	Ile	Thr	Ser
			210				215					220				
5	Arg	Asp	Ser	Arg	Gly	Asn	Asn	Ile	Leu	His	Ala	Leu	Val	Thr	Val	Ala
	225					230					235					240
	Glu	Asp	Phe	Lys	Thr	Gln	Asn	Asp	Phe	Val	Lys	Arg	Met	Tyr	Asp	Met
					245					250					255	
	Ile	Leu	Leu	Arg	Ser	Gly	Asn	Trp	Glu	Leu	Glu	Thr	Thr	Arg	Asn	Asn
10				260					265					270		
	Asp	Gly	Leu	Thr	Pro	Leu	Gln	Leu	Ala	Ala	Lys	Met	Gly	Lys	Ala	Glu
			275					280					285			
	Ile	Leu	Lys	Tyr	Ile	Leu	Ser	Arg	Glu	Ile	Lys	Glu	Lys	Arg	Leu	Arg
		290					295					300				
15	Ser	Leu	Ser	Arg	Lys	Phe	Thr	Asp	Trp	Ala	Tyr	Gly	Pro	Val	Ser	Ser
	305				310						315					320
	Ser	Leu	Tyr	Asp	Leu	Thr	Asn	Val	Asp	Thr	Thr	Thr	Asp	Asn	Ser	Val
				325					330					335		
	Leu	Glu	Ile	Thr	Val	Tyr	Asn	Thr	Asn	Ile	Asp	Asn	Arg	His	Glu	Met
20			340					345					350			
	Leu	Thr	Leu	Glu	Pro	Leu	His	Thr	Leu	Leu	His	Met	Lys	Trp	Lys	Lys
		355					360					365				
	Phe	Ala	Lys	His	Met	Phe	Phe	Leu	Ser	Phe	Cys	Phe	Tyr	Phe	Phe	Tyr
		370				375					380					
25	Asn	Ile	Thr	Leu	Thr	Leu	Val	Ser	Tyr	Tyr	Arg	Pro	Arg	Glu	Glu	Glu
	385				390						395					400
	Ala	Ile	Pro	His	Pro	Leu	Ala	Leu	Thr	His	Lys	Met	Gly	Trp	Leu	Gln
				405				410					415			
	Leu	Leu	Gly	Arg	Met	Phe	Val	Leu	Ile	Trp	Ala	Met	Cys	Ile	Ser	Val
30			420					425					430			
	Lys	Glu	Gly	Ile	Ala	Ile	Phe	Leu	Leu	Arg	Pro	Ser	Asp	Leu	Gln	Ser
		435					440					445				
	Ile	Leu	Ser	Asp	Ala	Trp	Phe	His	Phe	Val	Leu	Val	Pro	Arg	Leu	Pro
		450				455					460					
35	Arg	Ala	Gly	His	Gly	Pro	Gly	Leu	Gly	Glu	His	Ala	Leu	Leu	Tyr	Ala
	465				470					475						480
	Gly	Phe	Pro	Val	His	Gly	His	Val	Gln	Arg	His	Asp	Pro	Glu	Gly	His
				485				490					495			
40	Phe	Ala														

<210> 5

<211> 2373

<212> DNA



<213> Homo sapiens

<400> 5

	atgaaagccc accccaagga gatggtgcct ctcalgggca agagagttgc tgccccagct	60
5	gggaaccctg cegtccctgcc agagaagagg ccggcggaga tcacccccac aaagaagagt	120
	gcacacttct tcctggagat agaagggttt gaacccaacc ccacagttgc caagacctct	180
	cctcctgtct tctccaagcc catggattcc aacatccggc agtgcattct tggttaactgt	240
	gatgacatgg actccccca gtctcctcar gatgatgtga cagagacccc atccaatccc	300
	aacagcccca gtgcacagct ggccaaggaa gagcagagga ggaaaaagrg gcggctgaag	360
10	aagcgcattct ttgcagccgt gtctgagggc tgcgtggagg agttggtaga gttgctggtg	420
	gagctgcagg agctttgcag gcggcgccat gatgaggatg tgcttgactt cctcatgcac	480
	aagctgacgg cctccgacac ggggaagacc tgctgatga aggccttggt aaacatcaac	540
	cccaacacca aggagatmgt gcggatcctg cttgcctttg ctgaagagaa cgacatcctg	600
	ggcaggttca tcaacgccga gtacacagag gaggcctatg aagggcagac ggcgctgaac	660
15	atcgccatcg agcggcggca gggggacatc gcagccctgc tcatcgccgc cggcgccgac	720
	gtcaacgcgc acgccaaggg ggccctcttc aacccaagt accaacaaga aggcctctac	780
	ttcggtgaga cggccctggc cctggcagca tgcaccaacc agcccgagat tgtgcagctg	840
	ctgatggagc acgagcagac ggacatcacc tcgcgggact cacgaggcaa caacatcctt	900
	cacgccttgg tgaccgtggc cgaggacttc aagacrcaga atgactttgt gaagcgcattg	960
20	tacgacatga tcctactgcg gagtggcaac tgggagctgg agaccactcg caacaacgat	1020
	ggcctcacgc cgctgcagct ggccgccaag atgggcaagg cggagatcct gaagtacatc	1080
	ctcagtcgtg agatcaagga gaagcggctc cggagcctgt ccaggaagtt caccgactgg	1140
	gcgtacggac ccgtgtcatc ctccctctac gacctacca acgtggacac caccacggac	1200
	aactcagtgc tggaaatcac tgtctacaac accaacatcg acaaccggca tgagatgctg	1260
25	accctggagc cgctgcacac gctgctgcat atgaagtgga agaagtttgc caagcacatg	1320
	ttctttctgt ccttctgctt ttatttcttc tacaacatca ccctgaccct cgtctcgtac	1380
	taccgcccc gggaggagga ggccatcccg cacccttgg ccctgacgca caagatgggg	1440
	tggctgcagc tcctagggag gatgtttgtg ctcatctggg ccatgtgcat ctctgtgaaa	1500
	gagggcattg ccattctcct gctgagaccc tcggatctgc agtccatcct ctcgatgcc	1560
30	tggttccact ttgtcttttt tatccaagct gtgcttgtga tactgtctgt cttcttgtac	1620
	ttgtttgcct acaaagagta cctcgctgc ctctgtctgg ccatggccct gggctgggcg	1680
	aacatgctct actatacgcg gggtttccag tccatgggca tgtacagcgt catgatccag	1740
	aaggtyatth tgcatgatgt tctgaagttc ttgtttgat atatcgtgtt ttgtcttgga	1800
	tttgagtag ccttggcctc gctgatcgag aagtgtccca aagacaacaa ggactgcagc	1860
35	tcctacggca gcttcagyga cgagtgctg gaactcttca agctcaacct aggcctgggt	1920
	gayctgaaca tccagcagaa ctccaagtat cccattctct ttctgttctt gctcatcacc	1980
	tatgtcatcc tcacctttgt tctcctctc aacatgctca ttgctctgat gggcgagact	2040
	gtggagaacg tctccaagga gagcgaacgc atctggcgcc tgcagagagc caggaccatc	2100
	ttggagtthg agaaaatgth accagaatgg ctgaggagca gattccggat gggagagctg	2160
40	tgcaaagtgg ccgaggatga tttccgactg tgtttgcgga tcaatgaggt gaagtggact	2220
	gaatggaaga cgcacgtctc cttccttaac gaagaccgg ggctgtaag acgaacagat	2280
	ttcaacaaaa tccaagattc ttccaggaa aacagcaaaa cactctcaa tgcatttgaa	2340
	gaagtcgagg aattcccggg aacctcggtg tag	2373

<210> 6  
 <211> 790  
 <212> PRT  
 <213> Homo sapiens

5

<400> 6

	Met	Lys	Ala	His	Pro	Lys	Glu	Met	Val	Pro	Leu	Met	Gly	Lys	Arg	Val
	1				5				10					15		
	Ala	Ala	Pro	Ser	Gly	Asn	Pro	Ala	Val	Leu	Pro	Glu	Lys	Arg	Pro	Ala
10				20					25					30		
	Glu	Ile	Thr	Pro	Thr	Lys	Lys	Ser	Ala	His	Phe	Phe	Leu	Glu	Ile	Glu
			35					40					45			
	Gly	Phe	Glu	Pro	Asn	Pro	Thr	Val	Ala	Lys	Thr	Ser	Pro	Pro	Val	Phe
		50					55					60				
15	Ser	Lys	Pro	Met	Asp	Ser	Asn	Ile	Arg	Gln	Cys	Ile	Ser	Gly	Asn	Cys
	65					70					75				80	
	Asp	Asp	Met	Asp	Ser	Pro	Gln	Ser	Pro	Gln	Asp	Asp	Val	Thr	Glu	Thr
					85					90				95		
	Pro	Ser	Asn	Pro	Asn	Ser	Pro	Ser	Ala	Gln	Leu	Ala	Lys	Glu	Glu	Gln
20				100					105					110		
	Arg	Arg	Lys	Lys	Gly	Arg	Leu	Lys	Lys	Arg	Ile	Phe	Ala	Ala	Val	Ser
			115					120					125			
	Glu	Gly	Cys	Val	Glu	Glu	Leu	Val	Glu	Leu	Leu	Val	Glu	Leu	Gln	Glu
		130					135					140				
25	Leu	Cys	Arg	Arg	Arg	His	Asp	Glu	Asp	Val	Pro	Asp	Phe	Leu	Met	His
	145					150					155				160	
	Lys	Leu	Thr	Ala	Ser	Asp	Thr	Gly	Lys	Thr	Cys	Leu	Met	Lys	Ala	Leu
					165					170				175		
	Leu	Asn	Ile	Asn	Pro	Asn	Thr	Lys	Glu	Ile	Val	Arg	Ile	Leu	Leu	Ala
30				180					185					190		
	Phe	Ala	Glu	Glu	Asn	Asp	Ile	Leu	Gly	Arg	Phe	Ile	Asn	Ala	Glu	Tyr
			195					200					205			
	Thr	Glu	Glu	Ala	Tyr	Glu	Gly	Gln	Thr	Ala	Leu	Asn	Ile	Ala	Ile	Glu
		210					215					220				
35	Arg	Arg	Gln	Gly	Asp	Ile	Ala	Ala	Leu	Leu	Ile	Ala	Ala	Gly	Ala	Asp
	225					230					235				240	
	Val	Asn	Ala	His	Ala	Lys	Gly	Ala	Phe	Phe	Asn	Pro	Lys	Tyr	Gln	His
				245						250				255		
	Glu	Gly	Phe	Tyr	Phe	Gly	Glu	Thr	Pro	Leu	Ala	Leu	Ala	Ala	Cys	Thr
40				260					265					270		
	Asn	Gln	Pro	Glu	Ile	Val	Gln	Leu	Leu	Met	Glu	His	Glu	Gln	Thr	Asp
			275					280					285			
	Ile	Thr	Ser	Arg	Asp	Ser	Arg	Gly	Asn	Asn	Ile	Leu	His	Ala	Leu	Val
		290					295					300				

	Thr Val Ala Glu Asp Phe Lys Thr Gln Asn Asp Phe Val Lys Arg Met	
	305	310 315 320
	Tyr Asp Met Ile Leu Arg Ser Gly Asn Trp Glu Leu Glu Thr Thr	
		325 330 335
5	Arg Asn Asn Asp Gly Leu Thr Pro Leu Gln Leu Ala Ala Lys Met Gly	
		340 345 350
	Lys Ala Glu Ile Leu Lys Tyr Ile Leu Ser Arg Glu Ile Lys Glu Lys	
		355 360 365
	Arg Leu Arg Ser Leu Ser Arg Lys Phe Thr Asp Trp Ala Tyr Gly Pro	
10		370 375 380
	Val Ser Ser Ser Leu Tyr Asp Leu Thr Asn Val Asp Thr Thr Thr Asp	
		385 390 395 400
	Asn Ser Val Leu Glu Ile Thr Val Tyr Asn Thr Asn Ile Asp Asn Arg	
		405 410 415
15	His Glu Met Leu Thr Leu Glu Pro Leu His Thr Leu Leu His Met Lys	
		420 425 430
	Trp Lys Lys Phe Ala Lys His Met Phe Phe Leu Ser Phe Cys Phe Tyr	
		435 440 445
	Phe Phe Tyr Asn Ile Thr Leu Thr Leu Val Ser Tyr Tyr Arg Pro Arg	
20		450 455 460
	Glu Glu Glu Ala Ile Pro His Pro Leu Ala Leu Thr His Lys Met Gly	
		465 470 475 480
	Trp Leu Gln Leu Leu Gly Arg Met Phe Val Leu Ile Trp Ala Met Cys	
		485 490 495
25	Ile Ser Val Lys Glu Gly Ile Ala Ile Phe Leu Leu Arg Pro Ser Asp	
		500 505 510
	Leu Gln Ser Ile Leu Ser Asp Ala Trp Phe His Phe Val Phe Phe Ile	
		515 520 525
	Gln Ala Val Leu Val Ile Leu Ser Val Phe Leu Tyr Leu Phe Ala Tyr	
30		530 535 540
	Lys Glu Tyr Leu Ala Cys Leu Val Leu Ala Met Ala Leu Gly Trp Ala	
		545 550 555 560
	Asn Met Leu Tyr Tyr Thr Arg Gly Phe Gln Ser Met Gly Met Tyr Ser	
		565 570 575
35	Val Met Ile Gln Lys Val Ile Leu His Asp Val Leu Lys Phe Leu Phe	
		580 585 590
	Val Tyr Ile Val Phe Leu Leu Gly Phe Gly Val Ala Leu Ala Ser Leu	
		595 600 605
	Ile Glu Lys Cys Pro Lys Asp Asn Lys Asp Cys Ser Ser Tyr Gly Ser	
40		610 615 620
	Phe Ser Asp Ala Val Leu Glu Leu Phe Lys Leu Thr Ile Gly Leu Gly	
		625 630 635 640
	Asp Leu Asn Ile Gln Gln Asn Ser Lys Tyr Pro Ile Leu Phe Leu Phe	
		645 650 655

Leu Leu Ile Thr Tyr Val Ile Leu Thr Phe Val Leu Leu Leu Asn Met  
 660 665 670  
 Leu Ile Ala Leu Met Gly Glu Thr Val Glu Asn Val Ser Lys Glu Ser  
 675 680 685  
 5 Glu Arg Ile Trp Arg Leu Gln Arg Ala Arg Thr Ile Leu Glu Phe Glu  
 690 695 700  
 Lys Met Leu Pro Glu Trp Leu Arg Ser Arg Phe Arg Met Gly Glu Leu  
 705 710 715 720  
 10 Cys Lys Val Ala Glu Asp Asp Phe Arg Leu Cys Leu Arg Ile Asn Glu  
 725 730 735  
 Val Lys Trp Thr Glu Trp Lys Thr His Val Ser Phe Leu Asn Glu Asp  
 740 745 750  
 Pro Gly Pro Val Arg Arg Thr Asp Phe Asn Lys Ile Gln Asp Ser Ser  
 755 760 765  
 15 Arg Asn Asn Ser Lys Thr Thr Leu Asn Ala Phe Glu Glu Val Glu Glu  
 770 775 780  
 Phe Pro Glu Thr Ser Val  
 785 790  
 20 <210> 7  
 <211> 277  
 <212> DNA  
 <213> Homo sapiens  
 25 <400> 7  
 ttttaatctt gctaattaat tcttggaata atcaggaacg aaacagacaa ctttaagaaa 60  
 atattgttct tacttagact atactgaact gctatgtgcc ggtgaagaga agtytgtatg 120  
 ccagagcggc cgctgaattc tagaagccgt cctgccagag aagaggccgg cggagatcac 180  
 cccacaaaag aagagtgcac acttcttcct ggagatagaa gggtttgaac ccaacccac 240  
 30 agttgccaag acctctctc ctgtcttctc caagccc 277  
 35 <210> 8  
 <211> 2612  
 <212> DNA  
 <213> Homo sapiens  
 40 <400> 8  
 caggtggtc agccagttct gcctctgacg cctcattcca gccatccctc tgcttgaat 60  
 gagagcttcc cgccgcctca gccacagtcc caccggggg ccttggggcc cagacatgcg 120  
 gtgatctcag ggcaagggtt gccacgacca cccagaacct caccagccat gaaagccac 180  
 cccaaggaga tggtgctct catgggcaag agagttgctg ccccagtgga gaacctgcc 240  
 gtctgtccag agaagaggcc ggcgagatc acccccacaa agaagagtgc acacttcttc 300  
 ctggagatag aagggtttga acccaacccc acagttgcca agacctctcc tctgtcttc 360  
 tccaagccca tggattccaa catccggcag tgcattctctg gtaactgtga tgacatggac 420

	tccccccagt ctcctcarga tgatgtgaca gagaccccat ccaatcccaa cagccccagt	480
	gcacagctgg ccaaggaaga gcagaggagg aaaaagrggc ggctgaagaa ggcgcatctt	540
	gcagccgtgt ctgagggtg cgtggaggag ttggtagagt tgctgggtgga gctgcaggag	600
	ctttgcaggc ggcgccatga tgaggatgtg cctgacttcc tcatgcacaa gctgacggcc	660
5	tccgacacgg ggaagacctg cctgatgaag gccttgtaa acatcaaccc caacaccaag	720
	gagatmgtgc ggatcctgct tgcctttgct gaagagaacg acatcctggg caggttcatc	780
	aacgccgagt acacagagga ggcctatgaa gggcagacgg cgctgaacat cgccatcgag	840
	cggcggcagg gggacatcgc agccctgctc atcgccgccg gcgccgacgt caacgcgcac	900
	gccaaggggg ccttcttcaa ccccaagtac caacacgaag gcttctactt cgggtgagacg	960
10	cccctggccc tggcagcatg caccaaccag cccgagattg tgcagctgct gatggagcac	1020
	gagcagacgg acatcacctc gggggactca cgaggcaaca acatccttca cgccctggtg	1080
	accgtggccg aggacttcaa gacrcagaat gactttgtga agcgcatgta cgacatgac	1140
	ctactgcgga gtggcaactg ggagctggag accactcgca acaacgatgg cctcacgccg	1200
	ctgcagctgg ccgccaagat gggcaaggcg gagatcctga agtacatcct cagtcgtgag	1260
15	atcaaggaga agcggctccg gagcctgtcc aggaagtcca ccgactgggc gtacggaccc	1320
	gtgtcatcct ccctctacga cctcaccaac gtggacacca ccacggacaa ctcagtctg	1380
	gaaatcactg tctacaacac caacatcgac aaccggcatg agatgctgac cctggagccg	1440
	ctgcacacgc tgctgcatat gaagtggaa aagtttgcca agcacatgtt ctttctgtcc	1500
	ttctgctttt atttcttcta caacatcacc ctgaccctcg tctcgtacta ccgccccgg	1560
20	gaggaggagg ccatcccga ccccttggcc ctgacgcaca agatggggtg gctgcagctc	1620
	ctagggagga tgtttgtgct catctgggcc atgtgcatct ctgtgaaaga gggcattgcc	1680
	atcttctctg tgagaccctc ggatctgcag tccatcctct cggatgcctg gttccacttt	1740
	gtctttttta tccaagctgt gcttgtgata ctgtctgtct tcttgtactt gtttgctac	1800
	aaagagtacc tcgcctgcct cgtgctggcc atggccctgg gctgggcgaa catgctctac	1860
25	tatacgcggg gtttccagtc catgggcatg tacacgtca tgatccagaa ggtyattttg	1920
	catgatgttc tgaagttctt gtttgtatat atcgtgtttt tgcttgatt tggagtagcc	1980
	ttggcctcgc tgatcgagaa gtgtccaaa gacaacaagg actgcagctc ctacggcagc	2040
	ttcagygacg cagtgcctgga actcttcaag ctaccatag gcctgggtga yctgaacatc	2100
	cagcagaact ccaagtatcc cattctcttt ctgttctctg tcatcaccta tgtcatctc	2160
30	acctttgttc tctctctcaa catgctcatt gctctgatgg gcgagactgt ggagaacgtc	2220
	tccaaggaga gcgaacgcat ctggcgccctg cagagagcca ggaccatctt ggagtttgag	2280
	aaaatgttac cagaatggct gaggagcaga ttccgatgg gagagctgtg caaagtggcc	2340
	gaggatgatt tccgactgtg tttgcggatc aatgaggtga agtggactga atggaagacg	2400
	cacgtctcct tccttaacga agaccgggg cctgtaagac gaacagattt caacaaaatc	2460
35	caagattctt ccaggaacaa cagcaaaacc actctcaatg catttgaaga agtcgaggaa	2520
	ttcccggaaa cctcgggtga gaagcggaac ccagagctgg tgtgcgcgtg cgctgtctgg	2580
	cgetgcaggc ggagtcaccg actctgtgca ga	2612

	<210> 9
40	<211> 2286
	<212> DNA
	<213> Homo sapiens
	<400> 9

	atgaaagccc accccaagga gatggtgcct ctcatgggca agagagttgc tgccccagtc	60
	gggaaccctg ccgtcctgcc agagaagagg ccggcggaga tcacccccac aaagaagagt	120
	gcacacttct tcctggagat agaagggttt gaacccaacc ccacagttgc caagacctct	180
5	cctcctgtct tctccaagcc catggattcc aacatccggc agtgtgcaca gctggccaag	240
	gaagagcaga ggaggaaaaa grggcggctg aagaagcgca tctttgcagc cgtgtctgag	300
	ggctgcgtgg aggagttagt agagttgctg gtggagctgc aggagctttg caggcggcgc	360
	catgatgagg atgtgcctga cttcctcatg cacaagctga cggcctccga cacggggaag	420
	acctgcctga tgaaggcctt gttaaaccatc aacccaaca ccaaggagat mgtgcggatc	480
10	ctgcttgctt ttgctgaaga gaacgacatc ctgggcaggc tcatcaacgc cgagtacaca	540
	gaggaggcct atgaagggca gacggcgctg aacatcgcca tcgagcggcg gcagggggac	600
	atcgacagcc tgctcatcgc cgccggcgcc gacgtcaacg cgacgcgcaa gggggccttc	660
	ttcaacccca agtaccaaca cgaaggttc tacttcggtg agacgcccct ggccctggca	720
	gcatgcacca accagccga gattgtgcag ctgctgatgg agcacgagca gacggacatc	780
	acctgcggg actcacgagg caacaacatc cttcacgccc tggtgaccgt ggccgaggac	840
15	ttcaagacrc agaatgactt tgtgaagcgc atgtacgaca tgatcctact gcggagtggc	900
	aactgggagc tggagaccac tcgcaacaac gatggcctca cgccgctgca gctggccgcc	960
	aagatgggca aggcggagat cctgaagtac atcctcagtc gtgagatcaa ggagaagcgg	1020
	ctccggagcc tgtccaggaa gttcaccgac tgggcgtacg gaccctgtgc atcctccctc	1080
	tacgacctca ccaacgtgga caccaccacg gacaactcag tgctggaaat cactgtctac	1140
20	aacaccaaca tcgacaaccg gcatgagatg ctgaccctgg agccgctgca cacgctgctg	1200
	catatgaagt ggaagaagtt tgccaagcac atgttctttc tgtccttctg cttttatttc	1260
	ttctacaaca tcaccctgac cctcgtctcg tactaccgcc cccgggagga ggaggccatc	1320
	ccgcaccctt tggccctgac gcacaagatg ggggtggctgc agctcctagg gaggatgttt	1380
	gtgctcatct gggccatgtg catctctgtg aaagagggca ttgccatctt cctgctgaga	1440
25	ccctcggatc tgcagtccat cctctcggat gcctgggttc actttgtctt ttttatccaa	1500
	gctgtgcttg tgatactgtc tgtcttcttg tacttgtttg cctacaaaga gtacctgcc	1560
	tgccctcgtc tggccatggc cctgggctgg gcgaacatgc tctactatac gcggggtttc	1620
	cagtccatgg gcatgtacag cgtcatgac cagaaggtya ttttgcatga tgttctgaag	1680
	ttcttgtttg tatatatcgt gtttttgctt ggatttgag tagccttggc ctgctgac	1740
30	gagaagtgtc ccaaagacaa caaggactgc agctcctacg gcagcttcag ygacgcagtg	1800
	ctggaactct tcaagctcac cataggcctg ggtgayctga acatccagca gaactccaag	1860
	tatccattc tctttctgtt cctgctcatc acctatgtca tcctcacctt tgttctctc	1920
	ctcaacatgc tcattgctct gatgggcgag actgtggaga acgtctccaa ggagagcgaa	1980
	cgcactctggc gcctgcagag agccaggacc atcttgaggt ttgagaaaat gttaccagaa	2040
35	tggctgagga gcagattccg gatgggagag ctgtgcaaag tggccgagga tgatttccga	2100
	ctgtgtttgc ggatcaatga ggtgaagtgg actgaatgga agacgcacgt ctccttcctt	2160
	aacgaagacc cggggcctgt aagacgaaca gatttcaaca aaatccaaga ttcttccagg	2220
	aacaacagca aaaccactct caatgcattt gaagaagtcg aggaattccc ggaaacctcg	2280
40	gtgtag	2286

<210> 10  
 <211> 761  
 <212> PRT  
 <213> Homo sapiens

<400> 10

	Met	Lys	Ala	His	Pro	Lys	Glu	Met	Val	Pro	Leu	Met	Gly	Lys	Arg	Val
	1				5					10					15	
5	Ala	Ala	Pro	Ser	Gly	Asn	Pro	Ala	Val	Leu	Pro	Glu	Lys	Arg	Pro	Ala
					20				25					30		
	Glu	Ile	Thr	Pro	Thr	Lys	Lys	Ser	Ala	His	Phe	Phe	Leu	Glu	Ile	Glu
		35					40					45				
	Gly	Phe	Glu	Pro	Asn	Pro	Thr	Val	Ala	Lys	Thr	Ser	Pro	Pro	Val	Phe
10		50					55					60				
	Ser	Lys	Pro	Met	Asp	Ser	Asn	Ile	Arg	Gln	Cys	Ala	Gln	Leu	Ala	Lys
	65					70					75				80	
	Glu	Glu	Gln	Arg	Arg	Lys	Lys	Gly	Arg	Leu	Lys	Lys	Arg	Ile	Phe	Ala
					85				90					95		
15	Ala	Val	Ser	Glu	Gly	Cys	Val	Glu	Glu	Leu	Val	Glu	Leu	Leu	Val	Glu
				100				105					110			
	Leu	Gln	Glu	Leu	Cys	Arg	Arg	Arg	His	Asp	Glu	Asp	Val	Pro	Asp	Phe
		115					120					125				
	Leu	Met	His	Lys	Leu	Thr	Ala	Ser	Asp	Thr	Gly	Lys	Thr	Cys	Leu	Met
20		130					135					140				
	Lys	Ala	Leu	Leu	Asn	Ile	Asn	Pro	Asn	Thr	Lys	Glu	Ile	Val	Arg	Ile
	145				150						155				160	
	Leu	Leu	Ala	Phe	Ala	Glu	Glu	Asn	Asp	Ile	Leu	Gly	Arg	Phe	Ile	Asn
				165					170					175		
25	Ala	Glu	Tyr	Thr	Glu	Glu	Ala	Tyr	Glu	Gly	Gln	Thr	Ala	Leu	Asn	Ile
				180					185				190			
	Ala	Ile	Glu	Arg	Arg	Gln	Gly	Asp	Ile	Ala	Ala	Leu	Leu	Ile	Ala	Ala
		195					200					205				
	Gly	Ala	Asp	Val	Asn	Ala	His	Ala	Lys	Gly	Ala	Phe	Phe	Asn	Pro	Lys
30		210					215					220				
	Tyr	Gln	His	Glu	Gly	Phe	Tyr	Phe	Gly	Glu	Thr	Pro	Leu	Ala	Leu	Ala
	225				230						235			240		
	Ala	Cys	Thr	Asn	Gln	Pro	Glu	Ile	Val	Gln	Leu	Leu	Met	Glu	His	Glu
				245						250				255		
35	Gln	Thr	Asp	Ile	Thr	Ser	Arg	Asp	Ser	Arg	Gly	Asn	Asn	Ile	Leu	His
		260						265					270			
	Ala	Leu	Val	Thr	Val	Ala	Glu	Asp	Phe	Lys	Thr	Gln	Asn	Asp	Phe	Val
		275					280					285				
	Lys	Arg	Met	Tyr	Asp	Met	Ile	Leu	Leu	Arg	Ser	Gly	Asn	Trp	Glu	Leu
40		290					295					300				
	Glu	Thr	Thr	Arg	Asn	Asn	Asp	Gly	Leu	Thr	Pro	Leu	Gln	Leu	Ala	Ala
	305				310						315			320		
	Lys	Met	Gly	Lys	Ala	Glu	Ile	Leu	Lys	Tyr	Ile	Leu	Ser	Arg	Glu	Ile
				325						330				335		

	Lys	Glu	Lys	Arg	Leu	Arg	Ser	Leu	Ser	Arg	Lys	Phe	Thr	Asp	Trp	Ala	
					340				345					350			
	Tyr	Gly	Pro	Val	Ser	Ser	Ser	Leu	Tyr	Asp	Leu	Thr	Asn	Val	Asp	Thr	
			355					360					365				
5	Thr	Thr	Asp	Asn	Ser	Val	Leu	Glu	Ile	Thr	Val	Tyr	Asn	Thr	Asn	Ile	
		370					375					380					
	Asp	Asn	Arg	His	Glu	Met	Leu	Thr	Leu	Glu	Pro	Leu	His	Thr	Leu	Leu	
	385					390				395						400	
	His	Met	Lys	Trp	Lys	Lys	Phe	Ala	Lys	His	Met	Phe	Phe	Leu	Ser	Phe	
10					405					410					415		
	Cys	Phe	Tyr	Phe	Phe	Tyr	Asn	Ile	Thr	Leu	Thr	Leu	Val	Ser	Tyr	Tyr	
				420					425					430			
	Arg	Pro	Arg	Glu	Glu	Glu	Ala	Ile	Pro	His	Pro	Leu	Ala	Leu	Thr	His	
			435					440						445			
15	Lys	Met	Gly	Trp	Leu	Gln	Leu	Leu	Gly	Arg	Met	Phe	Val	Leu	Ile	Trp	
		450					455					460					
	Ala	Met	Cys	Ile	Ser	Val	Lys	Glu	Gly	Ile	Ala	Ile	Phe	Leu	Leu	Arg	
	465					470				475						480	
	Pro	Ser	Asp	Leu	Gln	Ser	Ile	Leu	Ser	Asp	Ala	Trp	Phe	His	Phe	Val	
20					485					490					495		
	Phe	Phe	Ile	Gln	Ala	Val	Leu	Val	Ile	Leu	Ser	Val	Phe	Leu	Tyr	Leu	
				500					505					510			
	Phe	Ala	Tyr	Lys	Glu	Tyr	Leu	Ala	Cys	Leu	Val	Leu	Ala	Met	Ala	Leu	
		515					520						525				
25	Gly	Trp	Ala	Asn	Met	Leu	Tyr	Tyr	Thr	Arg	Gly	Phe	Gln	Ser	Met	Gly	
		530					535					540					
	Met	Tyr	Ser	Val	Met	Ile	Gln	Lys	Val	Ile	Leu	His	Asp	Val	Leu	Lys	
	545					550				555						560	
	Phe	Leu	Phe	Val	Tyr	Ile	Val	Phe	Leu	Leu	Gly	Phe	Gly	Val	Ala	Leu	
30					565					570					575		
	Ala	Ser	Leu	Ile	Glu	Lys	Cys	Pro	Lys	Asp	Asn	Lys	Asp	Cys	Ser	Ser	
				580					585					590			
	Tyr	Gly	Ser	Phe	Ser	Asp	Ala	Val	Leu	Glu	Leu	Phe	Lys	Leu	Thr	Ile	
		595					600						605				
35	Gly	Leu	Gly	Asp	Leu	Asn	Ile	Gln	Gln	Asn	Ser	Lys	Tyr	Pro	Ile	Leu	
		610					615					620					
	Phe	Leu	Phe	Leu	Leu	Ile	Thr	Tyr	Val	Ile	Leu	Thr	Phe	Val	Leu	Leu	
	625					630				635						640	
	Leu	Asn	Met	Leu	Ile	Ala	Leu	Met	Gly	Glu	Thr	Val	Glu	Asn	Val	Ser	
40					645					650					655		
	Lys	Glu	Ser	Glu	Arg	Ile	Trp	Arg	Leu	Gln	Arg	Ala	Arg	Thr	Ile	Leu	
				660					665					670			
	Glu	Phe	Glu	Lys	Met	Leu	Pro	Glu	Trp	Leu	Arg	Ser	Arg	Phe	Arg	Met	
				675				680					685				



	Gly Glu Leu Cys Lys Val Ala Glu Asp Asp Phe Arg Leu Cys Leu Arg	
	690 695 700	
	Ile Asn Glu Val Lys Trp Thr Glu Trp Lys Thr His Val Ser Phe Leu	
	705 710 715 720	
5	Asn Glu Asp Pro Gly Pro Val Arg Arg Thr Asp Phe Asn Lys Ile Gln	
	725 730 735	
	Asp Ser Ser Arg Asn Asn Ser Lys Thr Thr Leu Asn Ala Phe Glu Glu	
	740 745 750	
10	Val Glu Glu Phe Pro Glu Thr Ser Val	
	755 760	
	<210> 11	
	<211> 22	
	<212> DNA	
15	<213> Artificial Sequence	
	<220>	
	<223> PCR primer	
20	<400> 11	
	cctcctcaac atgctcattg ct	22
	<210> 12	
	<211> 19	
25	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> PCR primer	
30	<400> 12	
	atgcgttcgc tctccttg	19
	<210> 13	
35	<211> 24	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
40	<223> Taqman probe	
	<400> 13	
	cggtctccac agtctcgccc atca	24

## Claims

1. An isolated polypeptide selected from the group consisting of:
  - (a) an isolated polypeptide encoded by a polynucleotide comprising the sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9;
  - (b) an isolated polypeptide comprising a polypeptide sequence having at least 95% identity to the polypeptide sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10;
  - (c) an isolated polypeptide having at least 95% identity to the polypeptide sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10; and
  - (d) fragments and variants of such polypeptides in (a) to (e).
2. The isolated polypeptide as claimed in claim 1 comprising the polypeptide sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10.
3. The isolated polypeptide as claimed in claim 1 which is the polypeptide sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10.
4. An isolated polynucleotide selected from the group consisting of:
  - (a) an isolated polynucleotide comprising a polynucleotide sequence having at least 95% identity to the polynucleotide sequence of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9;
  - (b) an isolated polynucleotide having at least 95% identity to the polynucleotide of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9;
  - (c) an isolated polynucleotide comprising a polynucleotide sequence encoding a polypeptide sequence having at least 95% identity to the polypeptide sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10;
  - (d) an isolated polynucleotide having a polynucleotide sequence encoding a polypeptide sequence having at least 95% identity to the polypeptide sequence of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10;
  - (e) an isolated polynucleotide with a nucleotide sequence of at least 100 nucleotides obtained by screening a library under stringent hybridization conditions with a labeled probe having the sequence of SEQ ID NO: 1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9 or a fragment thereof having at least 15 nucleotides;
  - (f) a polynucleotide which is the RNA equivalent of a polynucleotide of (a) to (e); or a polynucleotide sequence complementary to said isolated polynucleotide
- and polynucleotides that are variants and fragments of the above mentioned polynucleotides or that are complementary to above mentioned polynucleotides, over the entire length thereof.

5. An isolated polynucleotide as claimed in claim 4 selected from the group consisting of:
- (a) an isolated polynucleotide comprising the polynucleotide of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9;
  - 5 (b) the isolated polynucleotide of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5 or SEQ ID NO:9;
  - (c) an isolated polynucleotide comprising a polynucleotide sequence encoding the polypeptide of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10; and
  - (d) an isolated polynucleotide encoding the polypeptide of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6 or SEQ ID NO:10.
- 10
6. An expression system comprising a polynucleotide capable of producing a polypeptide of claim 1 when said expression vector is present in a compatible host cell.
7. A recombinant host cell comprising the expression vector of claim 6 or a membrane thereof
- 15 expressing the polypeptide of claim 1.
8. A process for producing a polypeptide of claim 1 comprising the step of culturing a host cell as defined in claim 7 under conditions sufficient for the production of said polypeptide and recovering the polypeptide from the culture medium.
- 20
9. An antibody immunospecific for the polypeptide of any one of claims 1 to 3.
10. A method for screening to identify compounds that stimulate or inhibit the function or level of the polypeptide of claim 1 comprising a method selected from the group consisting of:
- 25 (a) measuring or, detecting, quantitatively or qualitatively, the binding of a candidate compound to the polypeptide (or to the cells or membranes expressing the polypeptide) or a fusion protein thereof by means of a label directly or indirectly associated with the candidate compound;
  - (b) measuring the competition of binding of a candidate compound to the polypeptide (or to the cells or membranes expressing the polypeptide) or a fusion protein thereof in the presence of a
  - 30 labeled competitor;
  - (c) testing whether the candidate compound results in a signal generated by activation or inhibition of the polypeptide, using detection systems appropriate to the cells or cell membranes expressing the polypeptide;
  - (d) mixing a candidate compound with a solution containing a polypeptide of claim 1, to form a
  - 35 mixture, measuring activity of the polypeptide in the mixture, and comparing the activity of the mixture to a control mixture which contains no candidate compound; or

(e) detecting the effect of a candidate compound on the production of mRNA encoding said polypeptide or said polypeptide in cells, using for instance, an ELISA assay.



INVESTOR IN PEOPLE

**Application No:** GB 0126380.5  
**Claims searched:** All

**Examiner:** Dr Rowena Dinham  
**Date of search:** 1 July 2002

## **Patents Act 1977**

### **Search Report under Section 17**

#### **Databases searched:**

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.T):

Int Cl (Ed.7):

Other: ONLINE: WPI, EPODOC, JAPIO, MEDLINE, EMBASE, BIOSIS, SCISEARCH, CAPLUS, BLASTp, BLASTn

#### **Documents considered to be relevant:**

Category	Identity of document and relevant passage	Relevant to claims
X, E	WO 02/12340 A2 (INCYTE GENOMICS, INC.) See entire document especially SEQ ID No 1 and examples	1-10
X, E	WO 02/00722 A2 (MILLENNIUM PHARMACEUTICALS, INC.) See entire document especially SEQ ID No 3 and examples	1-10

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.